DENSITY FUNCTIONAL THEORY: VAN DER WAALS CORRECTIONS, SYMMETRY-SPHERICITY AND MACHINE LEARNING

A Dissertation submitted to
the Temple University Graduate Board

---

in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy (PhD)
in Physics

---

by
Shah Tanvir ur Rahman Chowdhury
August 2022

EXAMINING COMMITTEE MEMBERS:
Adrienn Ruzsinszky, Examining Committee Chair, Dept. of Physics
John P. Perdew, Advisory Committee Chair, Dept. of Physics
Qimin Yan, Examining Committee member, Dept. of Physics
Spiridoula Matsika, External member of Examining Committee, Dept. of Chemistry

This page intentionally left blank

ABSTRACT

van der Waals (vdW) or dispersion interaction dominates the weak bonding between the layers of a layered material or between a closed-shell molecule, monolayer, or multilayered material and a solid surface. The computation of materials properties, including physical adsorption of closed-shell species on solid surfaces, must include a description of the intermediate and long-range parts of the vdW interaction. Calculation of the vdW interaction by correlated wave-functions or random-phase-approximation (RPA) methods is not computationally efficient enough to describe the large systems or supercells that arise in many adsorption problems. Standard semilocal density functional approximations (e.g., PBE and SCAN) to the exchange-correlation energy are more practical, but the long-range part is not included in them. The main purpose of this part of the research, and its main accomplishment, was to develop a computationally efficient intermediate- and long-range correction to semi-local density functionals that would include the needed level of detail. We introduced and applied a new approach, the damped Zaremba-Kohn (dZK) method, to semiconducting substrates of finite thickness. The method has been applied to the adsorption of a graphene monolayer on bulk graphite, bulk hexagonal boron nitride, and multilayer transition metal dichalcogenides (one to four layers of MoS_2, WS_2, $MoSe_2$, and WSe_2). Furthermore, we extended the model to molecules adsorbed on a curved cylindrical conducting surface and combined this model with semilocal density functionals; Calculations were made for the molecules: ammonia (NH_3) and nitrogen dioxide (NO_2) at two adsorption sites, using the PBE, SCAN, PBE+dZK, and SCAN+dZK methods.

Turning now to the next part, we investigated several issues in density functional theory: the sensitivity to electron density of a hierarchy of nonempirical density functionals, the extent to which these density functionals approximate the exact functionals defined as constrained searches over wavefunctions versus ensembles, symmetry breaking and symmetry preservation, etc. We report results from several calculations with approximate density functionals which show that the total energies of non-spherical atoms are systematically lower than those for spherical atoms, a result which leads to appreciably improved molecular binding energies. In addition, we demonstrated that the energy consequences of the symmetry-breaking by self-consistent densities of open-shell atoms com-

puted with approximate functionals are small, justifying their use to compute atomization energies of molecules and solids, and justifying the use of the results to test the accuracy of the approximations.

Regarding the third part, our research presents a new approach to incorporate some exact constraints into machine learned density functionals (ML-DFT). ML-DFTs are one of the most exciting tools that have entered the material science toolbox in recent years. Recently, machine learning (ML) has been applied to parametrize exchange-correlation (XC) functionals without human domain knowledge by using kernel ridge regression (KRR), fully connected neural networks (NNs) and convolutional neural networks (CNNs). It is well-known that physical XC functionals must satisfy several exact conditions, such as coordinate scaling, spin scaling and derivative discontinuity. However, these exact conditions have not been incorporated implicitly into the machine learning modeling and preprocessing on large material datasets. In this work, we propose a schematic approach to incorporate a given physical constraint via contrastive learning. We then transfer the pretrained representation of electron density on an augmented molecular dataset which was generated by using the scaling property of exchange energy functionals based on the scaling factors chosen. The model with pretrained representation predicts exchange energies that satisfy the scaling relation, while the model trained on an unaugmented dataset gives poor predictions for the scaling-transformed electron density systems. Taken together, these findings suggest that pretraining task via contrastive learning can enhance the understanding of DFT theory for neural network models and generalize the application of NN-based XC functionals in a wide range of scenarios which are not always available experimentally but theoretically justified.

To my father Lutfor Rahman Shah, mother Tangila Rahman Chowdhury,
my sister Sadia Mustari Auntora, and wife Gabrielle Ford Chowdhury

ACKNOWLEDGEMENTS

First, I thank the Almighty God for His guidance and assistance.

I am extremely grateful to my supervisor, Prof. John P. Perdew for his invaluable advice, continuous support, and patience during my PhD study. I would also like to thank Dr. Hong Tang and Dr. Biswajit Santra for their technical support on my studies. Their immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life.

I thank Prof. Adrienn Ruzsinszky, Prof. Qimin Yan, and Weiyi Gong for their valuable help in my various research projects. I acknowledge Prof. Jim Napolitano, Prof. Andreas Metz, Prof. Peter Riseborough, and Prof. Charles J. Martoff for their constant support and encouragement. I am indebted to all my collaborators and funding agencies, without whom my research would not have been possible.

Finally, I would like to express my gratitude to my parents, and my sister. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my studies.

TABLE OF CONTENTS

This page intentionally left blank

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION TO DENSITY FUNCTIONAL THEORY

1.1 Introduction to Many-Body Problem

All materials that we encounter in our day-to-day life are composed of electrons and a few different kinds of nuclei which interact with each other electrostatically. Although electrons are much lighter than nuclei and require a quantum mechanical description, they usually move at a speed much less than that of light. Thus, in the absence of a time-dependent field, a system of M nuclei (at position $\mathbf{R}_1$, $\mathbf{R}_2$, ..., $\mathbf{R}_M$) and N electrons (at position $\mathbf{r}_1$, $\mathbf{r}_2$, ..., $\mathbf{r}_N$) can be solved using time-independent, non-relativistic Schrödinger wave equation (in atomic units) as [1],

$$\hat{H}\Psi(\mathbf{R}_1,\mathbf{R}_2,.....,\mathbf{R}_M,\mathbf{r}_1,\mathbf{r}_2,.....,\mathbf{r}_N) = E\Psi(\mathbf{R}_1,\mathbf{R}_2,.....,\mathbf{R}_M,\mathbf{r}_1,\mathbf{r}_2,.....,\mathbf{r}_N), \tag{1.1}$$

where $\Psi(\mathbf{R}_1,\mathbf{R}_2,.....,\mathbf{R}_M,\mathbf{r}_1,\mathbf{r}_2,.....,\mathbf{r}_N)$ is the many-body wave function associated with the hamiltonian ($\hat{H}$) which can be decomposed as

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ee} + \hat{V}_{ne}. \tag{1.2}$$

Here M_α is the ratio of mass of the α-th nucleus to that of an electron, and

$$\hat{T}_n = \sum_{\alpha=1}^{M} -\frac{1}{2M_\alpha}{\nabla_R}^2 \tag{1.3}$$

is the nuclear kinetic energy operator and

$$\hat{T}_e = \sum_{i=1}^{N} -\frac{1}{2}{\nabla_r}^2 \tag{1.4}$$

is the electronic kinetic energy operator. Similarly,

$$\hat{V}_{nn} = \sum_{\alpha=1}^{M} \sum_{\beta>\alpha}^{M} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \tag{1.5}$$

is the nucleus-nucleus interaction energy operator,

$$\hat{V}_{ee} = \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.6}$$

is the electron-electron interaction energy operator, and

$$\hat{V}_{ne} = -\sum_{\alpha=1}^{M} \sum_{i=1}^{N} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}_i|} \tag{1.7}$$

is the nucleus-electron interaction energy operator.

Although Eq (1.1-1) looks simple in form, it gives an exact solution only for hydrogen-like systems ($M = 1$ and $N = 1$), which is the simplest of two-body problems. When there are more than one electron and one nucleus in the system their motion is interconnected and almost impossible to decouple. Thus for many-body systems ($M + N > 2$), Eq (1.1-7) is utterly complicated to solve, and approximations are often required.

As the electrons are three orders of magnitude lighter than the nucleus, nuclear motion is much slower than that of electrons. In Born-Oppenheimer approximation [2], electrons can thus be assumed to drag along with nuclei with their motion parametrically dependent on nuclear positions but independent of their velocities. This allows us to assume *ansatz* of the form,

$$\Psi(\mathbf{R}, \mathbf{r}) = \phi_n(\mathbf{R})\phi_{el}(\mathbf{R}, \mathbf{r}), \tag{1.8}$$

where $\phi_n(\mathbf{R})$ denote the nuclear wave function and $\phi_{el}(\mathbf{R}, \mathbf{r})$ denote the electronic wave function. Thus for fixed nuclear positions,

$$\hat{H}_{el}\phi_{el}(\mathbf{R}, \mathbf{r}) = E_{el}\phi_{el}(\mathbf{R}, \mathbf{r}), \tag{1.9}$$

$$\hat{H}_{el} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}. \tag{1.10}$$

Since the nuclear positions are just the parameters for electronic wave function, $\hat{V}_{nn}$ shifts the energy eigenvalue by a constant, some consider including it in $\hat{H}_{el}$, as well. Although the use of the Born-Oppenheimer approximation simplifies the problem, the troublesome electron-electron interaction term makes it impossible to solve the electronic Hamiltonian exactly, reflecting the need for further approximations.

1.2 Wavefunction-Based Theory

Hartree proposed a self-consistent field approximation [3] in which an electron is assumed to move in an effective field due to all other $(N-1)$ electrons. Thus a many-body wavefunction could be written as a product of the N one-electron wavefunctions. However, Fock [4] soon noticed that such wavefunction is not antisymmetric and violates Pauli's principle. He proposed the many-body wavefunction to be the slater determinant of the N one-electron wave function. This way, it ensures antisymmetry in the many-body wavefunction. However, Hartree-Fock (HF) [3]-[5] approximation neglects the fact that the motion of the electrons is correlated as an electron tends to avoid other electrons. As its consequence, electrons are farther and lower in total energies than what HF theory predicts. There are various post-HF methods like Møller -Plesset perturbation theory [6], configuration interaction [7], coupled-cluster (CC) [8] etc., to incorporate the correlation effects. However, these methods are most feasible for atoms and molecules. These are too expensive for extended systems like solids.

1.3 Density functional Theory (DFT)

The solution of the electronic hamiltonian through convoluted many-body wavefunction can be avoided by using much simpler electron density in the framework of density functional theory. Electron density n($\mathbf{r}$) for a system of N electrons is defined as [1]:

$$n(\mathbf{r}) = N\sum_{\sigma}\sum_{\sigma_2}....\sum_{\sigma_N}\int d\mathbf{r}_2\int d\mathbf{r}_3....\int d\mathbf{r}_N|\Psi(\sigma,\mathbf{r},\sigma_2,\mathbf{r}_2,...,\sigma_N,\mathbf{r}_N)|^2 \tag{1.11}$$

where σ's are the spin indices and d$\mathbf{r}_N$ = d^3r_N. This immediately implies,

$$\int d\mathbf{r}\, n(\mathbf{r}) = N \tag{1.12}$$

We also define the spin densities as:

$$n_\sigma(\mathbf{r}) = N\sum_{\sigma_2}....\sum_{\sigma_N}\int d\mathbf{r}_2\int d\mathbf{r}_3....\int d\mathbf{r}_N|\Psi(\sigma,\mathbf{r},\sigma_2,\mathbf{r}_2,...,\sigma_N,\mathbf{r}_N)|^2 \tag{1.13}$$

1.3.1 Thomas-Fermi Model

Shortly after the introduction of Schrödinger's equation, Thomas [9] and Fermi [10] independently formulated what is known as the first density functional theory. In the Thomas-Fermi model, electrons are considered to be non-relativistic particles distributed uniformly in six-dimensional phase space. Their motion is assumed to govern "itself" by the effective

potential due to nuclear charge and the distribution of electrons. The total energy of the electrons in Thomas-Fermi theory is given by

$$E^{TF} = C\int d\mathbf{r}\ [n(\mathbf{r})]^{\frac{5}{3}} + \int d\mathbf{r}\ n(\mathbf{r})\ v(\mathbf{r}) + U[n] \tag{1.14}$$

where

$$C = \frac{3}{10}\ (3\pi^2)^{\frac{2}{3}}. \tag{1.15}$$

and,

$$U[n] = \frac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}^{'})}{|\mathbf{r}-\mathbf{r}^{'}|} d\mathbf{r}\ d\mathbf{r}^{'} \tag{1.16}$$

Thomas-Fermi (TF) theory is an important first step in the density functional theory formalism, however due to oversimplified assumptions it is not an accurate theory. First of all, the expression for kinetic energy (first term in RHS of Eq. (1.14)), is only approximate. Secondly, it doesn't account for the exchange energy and completely ignores the electron-electron correlation effects. As its consequences, TF theory can not bind atoms to form molecule [11]. Although imperfect, there were noteworthy efforts of Dirac and Weizsäcker to improve TF theory. The former made an effort to add the exchange energy term [12] while the latter proposed correction for the kinetic energy [13].

1.3.2 Hohenberg-Kohn (HK) Theory

In 1964, Hohenberg and Kohn gave an exact theory for the ground-state energy using electron density ($n(\mathbf{r})$ as the only variable) [14]. They introduced the universal functional $F[n]$, independent of the external potential, which plays a key role in determining the total ground-state energy [14].

The solution of the Schrödinger equation can be interpreted as a map between the set $\mathscr{V}$ of external potentials and the set $\mathscr{S}$ of resulting ground state. The sets $\mathscr{V}$ and $\mathscr{S}$ can be defined as [1],

$$\mathscr{S} = \{|\Psi_0\rangle\,|\,|\Psi_0\rangle \text{ ground state mapped onto one element of } \mathscr{V}\}$$

$$\mathscr{V} = \{v_{ext}|v_{ext} \text{ corresponding } |\Psi_0\rangle \text{ and } v'_{ext} \neq v_{ext} + \text{constant}\}$$

The Hohenberg-Kohn (HK) theorem can then be stated in the following fundamental statements [1].

1. There is an one-to-one [1] correspondence between the external potential v_{ext} in the Hamiltonian, the (non- degenerate) ground state $|\Psi_0\rangle$ resulting from solution of the Schrödinger equation and the associated ground state density n_0,

$$v_{ext} \Leftrightarrow |\Psi_0\rangle \Leftrightarrow n_0(\mathbf{r}) = \langle\Psi_0|\hat{n}(\mathbf{r})|\Psi_0\rangle$$

[1] A function $f : A \to B$ is called one-to-one if $a_1, a_2 \in A$ are such that $f(a_1) = f(a_2)$, then $a_1 = a_2$

Thus v_{ext}, $|\Psi_0\rangle$ and n_0 determine each other uniquely. In mathematical terms: the ground state is a unique functional of the ground state density, denoted as $|\Psi[n]\rangle$. Upon insertion of one element $n_0 \in \mathscr{N}$, this functional yields the ground state $|\Psi_0\rangle$ associated with this particular n_0,

$$|\Psi_0\rangle = |\Psi[n_0]\rangle$$

where,

$$\mathscr{N} = \{n_0 | n_0(\mathbf{r}) = \langle\Psi_0|\hat{n}(\mathbf{r})|\Psi_0\rangle\,, \quad |\Psi_0\rangle \in \mathscr{S}\}$$

It should be noted that no explicit information on v_{ext} is required to construct $|\Psi_0\rangle$ from n_0 : $|\Psi[n]\rangle$ has the same functional form for all kinds of many-particle systems with the same interaction. The same functional $|\Psi[n]\rangle$ applies to atoms, molecules and solids. The particular geometry of the systems under consideration is mediated by the structure of the density. $|\Psi[n]\rangle$ is therefore called universal. Clearly, the functional dependence of $|\Psi[n]\rangle$ on n must be extremely complicated.

2. The existence of the functional $|\Psi[n]\rangle$ leads to the statement that any ground state observable is a density functional,

$$O[n] = \langle\Psi[n]|\hat{O}|\Psi[n]\rangle$$

This is true, in particular, for the ground state energy, which turns out to be the most important density functional,

$$E[n] = \langle\Psi[n]|\hat{H}|\Psi[n]\rangle = F[n] + \int n(\mathbf{r}) v_{ext}(\mathbf{r}) d^3\mathbf{r} \tag{1.17}$$

and $F[n] = \langle\Psi[n]|\hat{T} + \hat{V_{ee}}|\Psi[n]\rangle$

The universal part $F[n]$ has been extracted from the total energy functional $E[n]$, in order to emphasize the fact that v_{ext} enters $E[n]$ at only one point.

3. There exists a minimum principle for $E[n]$: if n_0 is the ground state density corresponding to v_{ext}, one has for all densities,

$$E[n_0] < E[n_0'] \Leftrightarrow E_0 = \min\{E[n]\} \tag{1.18}$$

It is worthwhile to emphasize the restricted domain of the functional $E[n]$: only densities in N , i.e. ground state densities which are obtained by solution of the Schrödinger equation, are legitimate, however not arbitrary non-negative functions of $\mathbf{r}$ with finite norm.

1.3.3 v-representable and N-representable Electron Density

When the electron-density is associated with the antisymmetric ground-state wavefunction of a hamiltonian (having form of Eq. (1.10)) with external potential v($\mathbf{r}$), such density is called v-representable. The external potential need not necessarily be coulomb potential. However, both the HK theorems are based on the assumption that the densities are v-representable, there are densities in the single- and many-electrons systems that do not associate with the ground state wavefunction of the external potential [15, 16]. Fortunately, density functional theory can be formulated even if the densities satisfy a weaker condition called the N-representability condition. A density is N-representable if it is associated with some antisymmetric wave-function not necessarily the ground-state wavefunction of the external potential. Mathematically, they have following properties [17]:

$$n(\mathbf{r}) \geq 0, \qquad \int n(\mathbf{r}) d\mathbf{r} = N, \quad \text{and} \quad \int |\nabla (n(\mathbf{r}))^{\frac{1}{2}}|^2 \, d\mathbf{r} < \infty \, . \tag{1.19}$$

The minimum principle (Property 3 in section 1.3.2) indicates the possibility to determine the ground state density of a many-particle system by a variational equation [1],

$$\frac{\delta}{\delta n(\mathbf{r})} \left\{ E[n] - \mu \left(\int d^3\mathbf{r} \, n(\mathbf{r}) - N \right) \right\} \Bigg|_{n(\mathbf{r})=n_0(\mathbf{r})} = 0 \tag{1.20}$$

Unfortunately, the transition is not completely legitimate from a mathematical point of view [1]. The existence of the variational derivative $\frac{\delta E[n]}{\delta n(\mathbf{r})}$ requires the functional $E[n]$ to be defined on a sufficiently dense set of densities $n(\mathbf{r})$ (just as the differentiation of functions requires more than integer numbers as the domain of definition). If $E[n]$ is only defined for ground state densities resulting from the solution of the Schrödinger equation, we'll have to face the question whether there exists a potential v_{ext} for every given normalizable function $n(\mathbf{r})$, so that $n(\mathbf{r})$ is the ground state density corresponding to v_{ext}. If such a potential can be found the density is termed as v-representable.

Unfortunately, we can give explicit counterexamples, which demonstrate that the issue of v-representability is non-trivial for the functionals. The discussion of this problem is, however, mathematically somewhat involved.

The first counterexample is based on degeneracy. Assume that a ground energy level, for a given interacting or noninteracting H, has a degeneracy of order q. In that case, we can form the following ensemble density matrix [2], of rank q [18, 15],

$$D^{(N)}(1',2',...,N'|1,2,...,N) = \sum_{i=1}^{q} d_i \left|1',2',...,N'\right\rangle\left\langle 1,2,...,N\right| \tag{1.21}$$

[2] loosely following the notation of Ref. [15] Section V.

where $0 \leq d_i \leq 1$ and $\sum_{i=1}^{q} d_i = 1$ and the density is given by,

$$n(\mathbf{r}) = \sum_{i=1}^{q} d_i \, n_{ii}(\mathbf{r}) \tag{1.22}$$

Furthermore,

$$n_{ij}(\mathbf{r}) = N \int \phi_i(\mathbf{r}, \mathbf{r_2}, ..., \mathbf{r}_N)^* \phi_j(\mathbf{r}, \mathbf{r_2}, ..., \mathbf{r}_N) ds \, d\mathbf{r_2} \, ... \, d\mathbf{r}_N \tag{1.23}$$

where ϕ_i are the degenerate ground-state wave functions. We can, then, construct a statistical density matrix from these states and prove that the q unknown complex numbers a_i have to satisfy $2q^2$ linearly independent equations [15],

$$\sum_{i=1}^{q} \sum_{j=1}^{q} (\delta_{ij} d_j - a_i^* a_j) \, n_{ij}(\mathbf{r}) = 0 \tag{1.24}$$

Here, a_i's form the ground-state wave function $Q = \sum_{i=1}^{q} a_i \, \phi_i$ and $0 \leq |a_i| \leq 1$ [15].

Except when certain linear dependencies (excluding those due to $n_{ij}(\mathbf{r}) = n_{ji}(\mathbf{r})$) might happen to exist among the $n_{ij}(\mathbf{r})$, no solution of Eq. (1.24) for the a_i is generally possible for $q > 2$ [15], because the number of independent equations for the a_i; would exceed the number of unknowns [18]. Hence, for $q > 2$, wave function v-representability is not expected for those densities which can be constructed by the linear combinations given in Eq. (1.22). However, in Appendix A, we will show that for a special case, Levy's theorem is significantly true even when $n_{ij}(\mathbf{r}) = n_{ji}(\mathbf{r})$.

In fact, not every non-negative, normalizable function $n(\mathbf{r})$ is a reasonable candidate for which a ground state energy functional should be defined. Rather we require the components of the ground state energy to be finite separately [1],

$$\langle \Psi | \hat{T} | \Psi \rangle < \infty$$

$$\langle \Psi | \hat{v}_{ext} | \Psi \rangle < \infty$$

$$\langle \Psi | \hat{V_{ee}} | \Psi \rangle < \infty$$

The gradients of the admissible N-particle wavefunctions have to be square integrable. So, the first condition is satisfied.

$$T(\Psi) = N \sum_{\sigma_1, .., \sigma_N} \int d^3 r_1 .. d^3 r_N \, |\nabla_1 \Psi(r_1 \sigma_1, .., r_N \sigma_N)|^2 < \infty \tag{1.25}$$

In mathematical language, each component of $\nabla_i \Psi$ is in the set $\mathscr{L}^2\left(\mathbb{R}^{3N}\right)$ of square integrable functions over $\mathbb{R}^{3N}$ and each component of $\nabla n(r)^{1/2}$ is in the set $\mathscr{L}^2(\mathbb{R}^3)$ [1, 17]. Hence, both Ψ and $n(\mathbf{r})^{1/2}$ belong to Sobolev Space $\mathscr{H}^1(\mathbb{R}^{3N})$. Defined as [1],

$$\mathscr{H}^1(\mathbb{R}^{3N}) = \left\{ f | f \in \mathscr{L}^2\left(\mathbb{R}^{3N}\right), \nabla f \in \mathscr{L}^2\left(\mathbb{R}^{3N}\right) \right\} \tag{1.26}$$

However, $\Psi \in \mathscr{H}^1(\mathbb{R}^{3N})$ satisfies the constraint $\langle\Psi|\hat{V_{ee}}|\Psi\rangle < \infty$ [1]. On this basis, we can now consider the constraint $\langle\Psi|\hat{v}_{ext}|\Psi\rangle < \infty$, which leads to a characterization of the set of admissible external potentials. We first have to realize that any density in $\mathscr{S}$ also belongs to the larger space $\mathscr{L}^1(\mathbb{R}^3) \cap \mathscr{L}^2(\mathbb{R}^3)$, which is a Banach space (i.e. a complete normed vector space) [1, 17].

The integral $\langle\Psi|\hat{v}_{ext}|\Psi\rangle < \infty$ is thus finite, if,

$$v_{ext} \in \mathscr{L}^{3/2}(\mathbb{R}^3) + \mathscr{L}^{\infty}(\mathbb{R}^3)$$

where $\mathscr{L}^{\infty}(\mathbb{R}^3)$ denotes the space of all bounded functions.
Englisch *et. al.* [17] carried out a detailed study of all the factors that play a part in v-representability. They investigated some pathological examples of non-v-representable densities for single particles and argued convincingly that, the integral $\langle\Psi|\hat{v}_{ext}|\Psi\rangle > \infty$ is not finite, and,

$$v_{ext} \notin \mathscr{L}^{3/2}(\mathbb{R}^3) + \mathscr{L}^{\infty}(\mathbb{R}^3)$$

However, they put emphasis on the fact that the same proof is not applicable for interacting particles since,

$$n(\mathbf{r})/N = \int |\Psi(\mathbf{r}_1, .., \mathbf{r}_N)|^2 d^3\mathbf{r}_2 .. d^3\mathbf{r}_N$$

has infinitely many solutions [17] and it is not clear how to find the solution of,

$$\sum_i (\Delta_i \Psi)/\Psi - \sum \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} = v(\mathbf{r}_1, .., \mathbf{r}_N)$$

It is possible that for every Ψ, $\sum_i (\Delta_i \Psi)/\Psi$ is so singular for densities that the Hamiltonian $\Delta_i + \sum_i (\Delta_i \Psi)/\Psi$ is not a self-adjoint operator [17].

1.3.4 Kohn-Sham (KS) Formalism

The kinetic energy given by the TF model was approximate. Kohn and Sham introduced [19] the idea of orbitals in HK formalism to achieve accuracy for a major component of total energy, the kinetic energy, leaving the residual corrections to be handled separately. Consider the hamiltonian ($\hbar = m = 1$),

$$H_{eff} = \sum_{i=1}^{N} \left\{ -\frac{1}{2}\nabla_i^2 + v_{eff}(\mathbf{r}_i) \right\} \tag{1.27}$$

yielding the ground-state density equal to that of the corresponding interacting system. Here, $v_{eff}(\mathbf{r}_i)$ is the effective potential. The ground-state wavefunction of such system is exact and is given by

$$\phi_n = \frac{1}{\sqrt{N!}} \text{ det } [\phi_1 \phi_2 .... \phi_N \,], \tag{1.28}$$

where ϕ_i's are the Kohn-Sham orbitals obtained from the solution of one-electron hamiltonian as:

$$\left\{-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})\right\}\phi_i = \varepsilon_i\phi_i \tag{1.29}$$

where ε_i's are the Kohn-Sham orbital eigenvalues. These allow one to compute the non-interacting kinetic energy T_s as,

$$T_s[n] = -\sum_{i=1}^{N}\left\langle \phi_i\left|\frac{1}{2}\nabla^2\right|\phi_i\right\rangle \tag{1.30}$$

The universal functional $F[n]$ can be expressed in terms of T_s as,

$$F[n] = T_s[n] + U[n] + E_{xc}[n] \tag{1.31}$$

where, E_{xc} is the exchange-correlation energy. It contains the residual kinetic energy and the non-classical part of electron-electron interaction term. With these definitions, the minimization scheme can be written as,

$$0 = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta U[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = v_{eff}(\mathbf{r}) + \frac{\delta T_s[n]}{\delta n(\mathbf{r})} \tag{1.32}$$

where,

$$v_{eff}(\mathbf{r}) = v_{ext}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}(\mathbf{r}) = v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) \tag{1.33}$$

and,

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \tag{1.34}$$

is the exchange-correlation potential.

In Kohn-Sham scheme, one starts with a guess density to construct v_{eff} (Eq. 1.33). Kohn-Sham orbitals are then computed by solving N one-electron Schrödinger equations using that effective potential to compute the new density.

$$\left\{-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})\right\}\phi_i = \varepsilon_i\phi_i \tag{1.35}$$

$$n(\mathbf{r}) = \sum_{i=1}^{N}|\phi_i(\mathbf{r})|^2 \tag{1.36}$$

Thus the complicated 3N-dimensional many-body problem is reduced to a 3-dimensional one-body problem. Kohn and Sham employ a self-consistent approach to the problem by using a set of three equations (1.33), (1.35), and (1.36) known as the Kohn-Sham equations. We note here that the KS formalism can be generalized to include spin densities, as well. If $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$ are the ground-state spin-densities of the system of interacting electrons

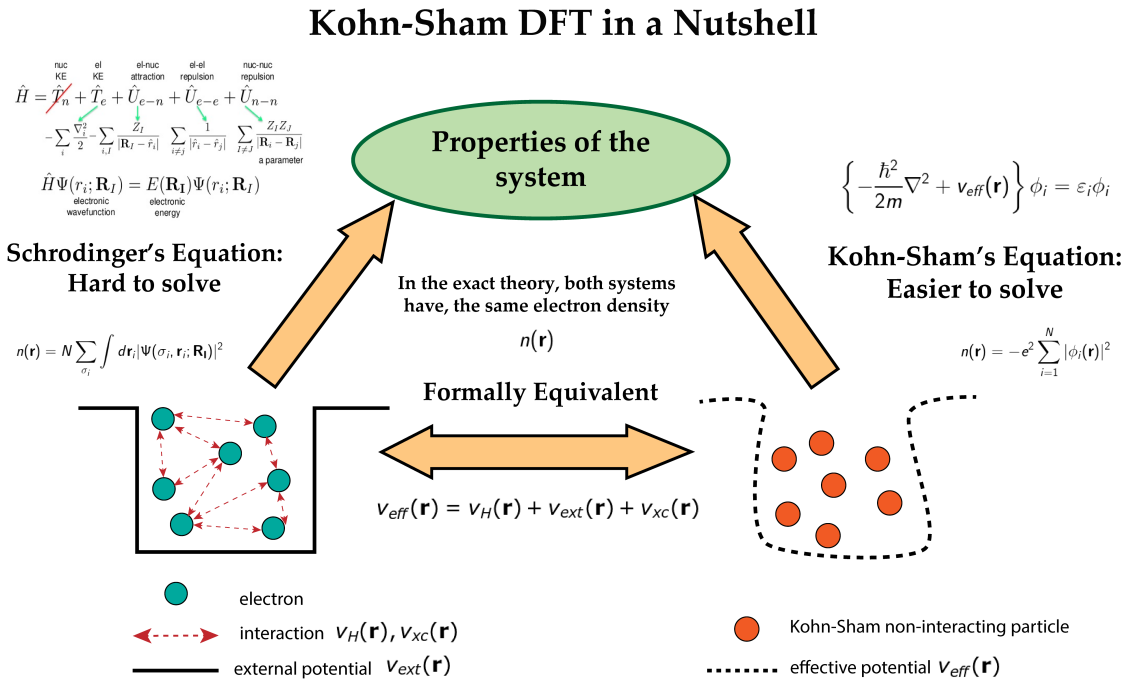

Figure 1.1: An overview of Kohn-Sham DFT

in the spin-dependent external potential $v^{\sigma}(\mathbf{r})$ ($\sigma = \uparrow, \downarrow$) then the set of KS equations are as follows:

$$\boxed{v^{\sigma}_{eff}(\mathbf{r}) = v^{\sigma}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v^{\sigma}_{xc}(\mathbf{r})} \tag{1.37}$$

$$\boxed{\left\{-\frac{1}{2}\nabla^2 + v^{\sigma}_{eff}(\mathbf{r})\right\}\phi_{i\sigma} = \varepsilon_i\,\phi_{i\sigma}} \tag{1.38}$$

$$\boxed{n_{\sigma}(\mathbf{r}) = \sum_{i=1}^{N} |\phi_{i\sigma}(\mathbf{r})|^2} \tag{1.39}$$

1.4 Exchange-Correlation Functional

In the previous section we introduced the exchange-correlation functional. It is expressed as a sum of the exchange functional (E_x) and the correlation functional (E_c) as,

$$E_{xc}[n] = E_x[n] + E_c[n]. \tag{1.40}$$

where,

$$E_x[n] = \langle \phi_n | \hat{V}_{ee} | \phi_n \rangle - U[n], \tag{1.41}$$

and

$$E_c[n] = \langle \Psi_n \mid \hat{T} + \hat{V}_{ee} \mid \Psi_n \rangle - \langle \Phi_n \mid \hat{T} + \hat{V}_{ee} \mid \Phi_n \rangle \ . \tag{1.42}$$

Kohn-Sham wavefunction (ϕ_n) yields a density $n(\mathbf{r})$ that minimizes the kinetic energy, which in fact is exact for such non-interacting fictitious system.

$$T_s[n] = \langle \phi_n \mid \hat{T} \mid \phi_n \rangle \tag{1.43}$$

However, the kinetic energy of the corresponding physical system is,

$$T[n] = \langle \Psi_n \mid \hat{T} \mid \Psi_n \rangle \tag{1.44}$$

We can re-write Eq (1.42) as:

$$E_c[n] = T_c[n] + U_c[n] \tag{1.45}$$

where,

$$T_c[n] = T[n] - T_s[n] \tag{1.46}$$

is the kinetic contribution to the correlation energy and

$$U_c[n] = \langle \Psi_n \mid \hat{V}_{ee} \mid \Psi_n \rangle - \langle \phi_n \mid \hat{V}_{ee} \mid \phi_n \rangle \tag{1.47}$$

is the potential contribution to the correlation energy.

We immediately see from the definition that $E_c[n]$ and $U_c[n]$ are non-positive while $T_c[n]$ is non-negative. From Eq. (1.42), $E_c[n] = 0$ for all one-electron systems. For practical purposes, E_x and E_c are approximated because their exact form, even if we had one, is computationally not feasible. However, there are constraints (some listed below) which an exact exchange and correlation functional shall obey. In addition, there are systems like uniform electron gas, hydrogen atom, jellium surface etc., where the density functionals can be exact or extremely accurate. These provide guidelines for constructing the approximate functionals.

1.4.1 Some Exact Constraints Satisfied by Density Functionals

1. **Negativity (trivial)** [20]:

$$E_x[n_\uparrow, n_\downarrow] < 0, \qquad E_c[n_\uparrow, n_\downarrow] \leq 0 \tag{1.48}$$

2. **Scaling Properties:** Under the uniform coordinate scaling of the density $n_\gamma(\mathbf{r}) \equiv \gamma^3 n(\gamma \mathbf{r})$, exchange and correlation functional satisfy the following [21, 22]:

$$E_x[n_\gamma] = \gamma E_x[n] \tag{1.49}$$

$$\lim_{\gamma\to\infty} E_c[n_\gamma] = constant \tag{1.50}$$

$$E_c[n_\gamma] > \gamma E_c[n], \qquad (\gamma > 1) \tag{1.51}$$

$$E_c[n_\gamma] < \gamma E_c[n], \qquad (\gamma < 1) \tag{1.52}$$

3. **Spin-scaling Relations** [23]:

$$E_x[n_\uparrow, n_\downarrow] = \frac{1}{2}E_x[2n_\uparrow] + \frac{1}{2}E_x[2n_\downarrow] \tag{1.53}$$

4. **Size-consistency** [24]: For well-separated densities n_A and n_B,

$$E_{xc}[n_A + n_B] = E_{xc}[n_A] + E_{xc}[n_B] \tag{1.54}$$

5. **Lieb and Oxford bound** [25]:

$$E_{xc}[n] \geq -1.679 \int d\mathbf{r}\, n^{\frac{4}{3}}(\mathbf{r}) \tag{1.55}$$

6. **Tight bound for two-electron densities** [25]:

$$E_x[n_2] > 1.174\, E_x^{LDA}[n_2] \tag{1.56}$$

Bounds on the integrated energy are imposed by bounding the energy density.

1.4.2 Local Density Approximation

Local density approximation (LDA) is the simplest exchange-correlation functional [19] that depends only on the electron density at each point in space. Here exchange and correlation energy are evaluated, assuming that the electron density is locally uniform. The expression for the exchange energy in LDA is

$$E_x^{LDA}[n(\mathbf{r})] = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_x^{UEG}(n(\mathbf{r})) \tag{1.57}$$

where,

$$\varepsilon_x^{UEG}(n(\mathbf{r})) = -\frac{3}{4\pi}(3\pi^2 n)^{\frac{1}{3}} \tag{1.58}$$

is the exact expression derived by Dirac [12] for the exchange energy per particle of uniform electron gas (UEG). Unlike exchange, correlation energy,

$$E_c^{LDA}[n(\mathbf{r})] = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_c^{UEG}(n(\mathbf{r}))\ , \tag{1.59}$$

is only known exactly for two extreme limits of the density [26, 27]. At various points between these limits, Ceperley-Alder [28] have computed correlation energy per particle

using Quantum Monte Carlo calculations. Due to lack of exact analytical form for correlation energy, there are a number of parameterizations [29, 30] that rely on interpolation and extrapolation techniques to fit those available data. Details on the accuracy of these parameterizations can be found in Ref [31]. For spin-dependent systems like open-shell atoms, magnetic materials etc., spin-polarized version of LDA (i.e. LSDA) is required. For LSDA, one defines the relative spin polarization

$$\zeta = \frac{n_\uparrow - n_\downarrow}{n} \tag{1.60}$$

where, $n = n_\uparrow + n_\downarrow$. ζ lies between 0 (spin unpolarized) and 1 (fully spin polarized). LSDA exchange and correlation energies are then computed for UEG, using the respective spin densities. LSDA by-design is exact for uniform electron density and is expected to be accurate for very-slowly varying densities. Unlike, TF theory, LSDA binds atoms to form molecule and predicts reasonable lattice constants in solids.

1.4.3 Generalized Gradient Approximation

LSDA only considers a local uniform density at a point in space. Therefore, they fail to capture the inhomogeneity required to describe systems in reality. The electron density, along with its gradient, can incorporate such inhomogeneity to some extent. Such a class of approximate functionals is called the generalized gradient approximation (GGA). Compared to LSDA, GGA's tend to improve total energies, atomization energies, etc. The exchange-correlation energy in GGAs is written as

$$\boxed{E_{xc}^{GGA}[n_\uparrow, n_\downarrow] = \int d\mathbf{r} f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow)} \tag{1.61}$$

GGA is constructed in such a way that it recovers the second-order gradient expansion for slowly-varying densities and is exact for the case of uniform density. There are a number of approximations [32, 33] for exchange and correlation energies at the GGA level [34]. A highly successful Perdew-Burke-Ernzerhof (PBE) [35] exchange-correlation functional combines PBE-exchange with PBE-correlation. It is a general-purpose GGA designed to satisfy 11 exact constraints. The GGA exchange-correlation energy is expressed in terms of the enhancement factor ($F_{xc}(s)$) over the local exchange (ε_x^{UEG}) as

$$E_{xc}^{GGA}[n_\uparrow, n_\downarrow] = \int d\mathbf{r} n(\mathbf{r}) \varepsilon_x^{UEG} F_{xc}(r_s, \zeta, s) \tag{1.62}$$

where $r_s = (3/4\pi n)^{\frac{1}{3}}$ and $s = |\nabla n|/2(3\pi^2)^{\frac{1}{3}} n^{\frac{4}{3}}$ is the dimensionless density gradient. GGAs differ from one-another in the parameterized-analytical form of $F_{xc}(s)$. LSDA is a special case of such enhancement factor where $s = 0$. Although, GGAs are superior over LSDA for

predicting total energies and atomization energies, they mostly overestimate bond lengths. In addition, they underestimate the band-gap of semiconductors and insulators [36, 37] and fail to sufficiently account for the dispersion interactions.

1.4.4 Meta-Generalized Gradient Approximation

An approximate functional can satisfy more exact constraints when the kinetic energy density of the Kohn-sham orbitals is used as an ingredient in addition to the electron density and its gradient. Such a class of approximate functionals is known as meta-generalized gradient approximations (MGGA). The exchange-correlation energy in MGGA's is

$$\boxed{E_{xc}^{MGGA}[n_\uparrow, n_\downarrow] = \int d\mathbf{r}\, n(\mathbf{r})\varepsilon_{xc}(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow, \tau_\uparrow, \tau_\downarrow)} \tag{1.63}$$

Although there are a number of MGGAs introduced till date [38]-[39], in the present assessment we focus our study mainly on the strongly constrained and appropriately normed (SCAN) [40] and revSCAN [41] MGGAs. SCAN satisfies all 17 known exact constraints that a MGGA can satisfy. Here, τ appears in the exchange-correlation energy density as a dimensionless variable

$$\alpha(\mathbf{r}) = \frac{\tau(\mathbf{r}) - \tau^w(\mathbf{r})}{\tau^{unif}(\mathbf{r})} \tag{1.64}$$

where

$$\tau^w(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|^2}{8n(\mathbf{r})} \tag{1.65}$$

is the von Weizsäcker kinetic energy density (single-density limit of τ) and

$$\tau^{unif}(\mathbf{r}) = \frac{3}{10}(3\pi^2)^{\frac{2}{3}} n(\mathbf{r})^{\frac{5}{3}} \tag{1.66}$$

is the uniform-density limit. The iso-orbital indicator (α) can recognize different bonding situations like covalent single ($\alpha = 0$), metallic ($\alpha = 1$) and weak ($\alpha >> 1$) bonding.

Some modifications to SCAN have been proposed to improve its accuracy for specific systems. The revSCAN [41] functional is a simple modification to the slowly varying limit of SCAN's correlation energy to eliminate the fourth-order term in SCAN's correlation energy density-gradient expansion. The TASK functional is a complete revision of SCAN designed to accurately predict band gaps while retaining the exact constraints placed on the exchange energy [42], though TASK uses a local spin-density approximation (LSDA) to model correlation. It is not expected that these modifications address the numerical inefficiencies of the parent functional. In recent work, Bartok and Yates propose a regularized SCAN termed "rSCAN" that aims to control SCAN's numerical challenges while changing as little as possible from the parent functional [43]. The resulting functional

shows greatly improved numerical stability and enables pseudopotential generation. While initial testing suggested that rSCAN maintained the accuracy and transferability of SCAN, expanded testing by Ref. [44, 45] shows that some transferability is lost, with accuracy for atomization energies particularly degraded [46].

Ref. [47] provides evidence to support the claim that rSCAN makes wide-ranging sacrifices in exact constraint adherence in order to make a numerically efficient meta-GGA. Ref. [47] argues convincingly that such sacrifices are needless and derives revisions to the rSCAN functional to restore exact constraint adherence without harming numerical efficiency. Ref. [47], then, applies these revisions to build a regularized restored SCAN functional, r^2SCAN, which recovers the most important exact constraints of SCAN (see Figure 1 of Ref. [47], for example). Finally, in one-electron systems, the semilocal r^2SCAN makes the functional converge much better on sparse grids, and even makes it a little more accurate.

1.4.5 Inadequacy of Semilocal Functionals

GGAs and MGGAs are combinedly called semilocal functionals. These are widely applied class of functionals for electronic structure calculations in physics and chemistry [48]. Recently, semilocal density functional theory calculations are attracting more interest of material scientists to search and predict materials of desired properties [49, 50]. Although semilocal functionals systematically improve upon the local spin density approximations, they are not perfect. Firstly, the semilocal self-exchange-correlation term can not perfectly cancel the nonlocal self-Hartree energy term leading to the spurious self-interaction error. It is more problematic for the description of the system in a non-equilibrium state like stretched bonds. Secondly, the semilocal exchange-correlation functionals can not fully account for the long-range electron-electron correlation effects. It is problematic for molecular dimers, molecular crystals, and interfaces where dispersion interaction is significant.

1.4.6 Hybrid Functionals

Mixing a portion of exact exchange (or HF exchange) [3, 4] with the (semi)local exchange and correlation improves the description of band gaps and molecular properties like atomization energies, bond-lengths, and vibrational frequencies [35]. Such a class of functionals is known as hyperGGAs or hybrid functionals. Here the exchange-correlation energy is

$$\boxed{E_{xc}^{hybrid} = E_x^{DFT} + a(E_x^{exact} - E_x^{DFT}) + E_c^{DFT}} \tag{1.67}$$

where a is the mixing parameter. Some of the popular hybrid functionals are PBE0 [51] which combines PBE exchange and correlation with the HF-exchange with $a = 0.25$, range-separated hybrid HSE06 [52] with $a = 0.2$ and a three-parameter B3LYP [33, 34] functional. In addition to improvements over the semilocal functionals for different properties, hybrid functionals also incorporate nonlocal effects which help reduce self-interaction errors in (semi)local functionals but do not completely eliminate them.

1.5 Nonlocal Correlation Functionals

The nonlocal correlation functionals combine with the semilocal density functional approximations (DFAs) such that the exchange-correlation energy can be expressed as,

$$E_{xc} = E_{xc}^{DFA} + E_c^{nl}, \tag{1.68}$$

where, E_{xc}^{DFA} is the contribution from the semilocal approximations, and [53],

$$\boxed{E_c^{nl} = \frac{1}{2}\int d\mathbf{r} \int d\mathbf{r}'\, n(\mathbf{r})\Phi[n](\mathbf{r},\mathbf{r}')\, n(\mathbf{r}')} \tag{1.69}$$

is the contribution from the nonlocal correlation functional. The nonlocal contribution comes from the so-called non-local kernel $\Phi[n](\mathbf{r},\mathbf{r}')$, which is a function of $\mathbf{r}-\mathbf{r}'$, electron density $n(\mathbf{r})$ and its gradient meaning the dispersion interaction is addressed in a general and seamless manner without explicitly depending on the orbitals. It can be shown that the non-local kernel $\Phi[n](\mathbf{r},\mathbf{r}')$ depends on the local effective frequency $\omega[n]$ and the range-separating function $f[n]$, where the differences between distinct nonlocal functionals reside in the specific functional forms of $\omega[n]$ and $f[n]$ [53]. We discuss, in brief, below a couple of popular nonlocal correlation functionals. Later, in chapter 2 and 3, we will take a closer look at them.

1.5.1 vdW-DFs

Dion *et al.* [54] introduced the first vdW-DFs (density functionals). They incorporate the vdW interactions seamlessly and ensure no corrections in the case of uniform electron density. Here the nonlocal correlations are expressed in the non-local kernel $\Phi[n](\mathbf{r},\mathbf{r}')$ in terms of the density-density interaction. They utilized revPBE [55] functional for semilocal exchange because, unlike most GGAs, it is not a dominant contributor of binding energies in rare-gas dimers. Although this vdW-DF showed some promising results when applied to gas-dimers and benzene dimers, the errors were large for molecular complexes. Lee *et al.* [56] revised the vdW-DF by revising the kernel and replacing revPBE by a more accurate

exchange functional revised PW86 [57] (rPW86 [58]). This so-called vdW-dF2 functional predicted the interaction energies and distances of the molecular dimers in the S22 dataset [59] very close to the available reference values. However, Yildrim *et al.* [60] showed that both these vdW-DFs significantly overbind adsorption of organic molecule benzene over transition metals unless paired with "less repulsive" [61] semilocal approximations like optPBE and optB88.

1.5.2 VV10 and its Revised Version (rVV10)

Vydrov *et al.* [62] introduced a useful nonlocal correlation functional that had similar form of E_c^{nl} as that for vdW-DFs. The total VV10 correlation energy is written as [62],

$$E_c^{VV10} = E_c^{nl} + \beta N \tag{1.70}$$

where,

$$\beta = \frac{1}{32}\left(\frac{3}{b}\right)^{\frac{3}{4}} \tag{1.71}$$

is the density-independent constant to ensure no corrections are made by the nonlocal functional in the case of uniform density. Here the nonlocality is addressed by the non-local kernel $\Phi[n](\mathbf{r},\mathbf{r}')$ and it has two adjustable parameters namely C and b which respectively control the asymptotic limit and short-range damping. The total exchange-correlation energy is thus written as [53],

$$E_{xc}^{VV10} = E_x^{rPW86} + E_c^{PBE} + E_c^{VV10}. \tag{1.72}$$

The original VV10 determined $b = 5.9$ and $C = 0.0093$ by fitting to the reference interaction energies of molecular dimers in the S22 dataset [59]. Sabatini *et al.* [63] modified VV10 such that the nonlocal correlation energy and its derivatives would be efficiently evaluated in the plane-wave framework as established by Roḿan, Pérez and Soler (RPS) [64]. Although the C parameter was the same, there was a slight change in the optimized b parameter by the new implementation ($b = 6.3$). S22 is the dataset [59] of diversely-bonded molecular dimers. So it is often used for determining the adjustable parameters. However, it is common, as well, to parameterize VV10 or rVV10 using simpler systems like rare-gas dimers and layered materials.

1.6 Adiabatic Connection (AC)

The ground-state total energy of an interacting many-body Hamiltonian can formally be obtained via the AC technique, in which a continuous set of coupling-strength λ dependent

Hamiltonians is introduced,

$$\hat{H}(\lambda) = \hat{H}_0 + \lambda \hat{H}_1(\lambda)$$

that "connect" a reference Hamiltonian $\hat{H}_0 = \hat{H}(\lambda = 0)$ with the target many-body Hamiltonian $\hat{H}$. While $\lambda = 1$ being the physical external potential of the fully-interacting system. The reference Hamiltonian $\hat{H}_0$, for $\lambda = 0$, is of the mean-field (MF) type, i.e., a simple summation over single-particle Hamiltonians. In the AC construction of the total energy, we introduce the ground-state wave function $|\Psi_\lambda\rangle$ for the λ-dependent, system such that

$$H(\lambda)\,|\Psi_\lambda\rangle = E(\lambda)\,|\Psi_\lambda\rangle$$

The interacting ground-state total energy can then be obtained with the aid of the Hellmann-Feynman theorem,

$$E(\lambda = 1) = E_0 + \int_0^1 d\lambda \left\langle \Psi_\lambda \middle| \left(H_1 + \lambda \frac{dH_1(\lambda)}{d\lambda} \right) \middle| \Psi_\lambda \right\rangle \tag{1.73}$$

where E_0 is the zeroth-order energy $\langle \Psi_\lambda | H_0 | \Psi_\lambda \rangle$. We note that the choice of the adiabatic-connection path is not unique. In DFT, the path is chosen such that the electron density is kept fixed at its physical value along the way.

1.6.1 Random Phase Approximation (RPA)

In KS-DFT the AC path is chosen such that the electron density is kept fixed. Eq. 1.73 can then be written as,

$$E_{xc} \sim \int d\lambda \int d^3\mathbf{r} \int d^3\mathbf{r}' \frac{n(\mathbf{r}) n_{xc}^{\lambda}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{1.74}$$

where,

$$n_{xc}^{\lambda}(\mathbf{r}, \mathbf{r}') = \frac{R(\hat{n}(\mathbf{r}), \hat{n}(\mathbf{r}'))}{n(\mathbf{r})} - \delta(\mathbf{r} - \mathbf{r}') \tag{1.75}$$

where, $R(\hat{n}(\mathbf{r}), \hat{n}(\mathbf{r}'))$ is the density-density correlations (fluctuations). Hence, $R(\hat{n}(\mathbf{r}), \hat{n}(\mathbf{r}')) = \langle \Psi_\lambda | \delta\hat{n}(\mathbf{r}) \delta\hat{n}(\mathbf{r}') | \Psi_\lambda \rangle$ with $\delta\hat{n}(\mathbf{r}) = \hat{n}(\mathbf{r}) - n(\mathbf{r})$ being the fluctuation of the density operator $\hat{n}(\mathbf{r})$ around its expectation value $n(\mathbf{r})$. It shows that the XC hole is related to the density-density correlation function. In physical terms, it describes how the presence of an electron at point $\mathbf{r}$ depletes the density of all other electrons at another point $\mathbf{r}'$

In the second step, the density-density correlations (fluctuations) $R(\hat{n}(\mathbf{r}), \hat{n}(\mathbf{r}'))$ are linked to the response properties (dissipation) of the system through the zero-temperature fluctuation-dissipation theorem [65],

$$R(\hat{n}(\mathbf{r}), \hat{n}(\mathbf{r}')) = -\frac{1}{\pi} \int_0^{\infty} d\omega \mathrm{Im}[\chi^{\lambda}(\mathbf{r}, \mathbf{r}', \omega)] \tag{1.76}$$

where $\chi^{\lambda}(\mathbf{r},\mathbf{r}',\omega)$ is the linear density response function of the (λ-scaled) system. We arrive at the renowned ACFD expression for the XC energy in DFT.

The reason that the above frequency integration can be performed along the imaginary axis originates from the analytical structure of $\chi^{\lambda}(\mathbf{r},\mathbf{r}',\omega)$ and the fact that it becomes real on the imaginary axis. The ACFD expression transforms the problem of computing the XC energy to one of computing the response functions of a series of fictitious systems along the AC path, which in practice have to be approximated as well. In this context the random-phase approximation is a particularly simple approximation of the response function.

$$\chi^{\lambda}_{RPA}(\mathbf{r},\mathbf{r}',\omega) = \chi^{0}(\mathbf{r},\mathbf{r}',\omega) + \int d\mathbf{r_1} d\mathbf{r_2}\, \chi^{0}(\mathbf{r},\mathbf{r_1},\omega)\, \frac{\lambda}{|\mathbf{r_1}-\mathbf{r_2}|}\, \chi^{\lambda}(\mathbf{r_2},\mathbf{r}',\omega) \tag{1.77}$$

$\chi^{0}(\mathbf{r},\mathbf{r}',\omega)$ is the independent-particle response function of the KS reference system at $\lambda = 0$ and is known explicitly in terms of the single-particle KS orbitals $\phi_i(\mathbf{r})$.

Thus, the XC energy in RPA can be separated into an exact exchange and the RPA correlation term,

$$\boxed{E^{RPA}_{xc} = E^{exact}_{x} + E^{RPA}_{c}} \tag{1.78}$$

where,

$$E^{exact}_{x} = -\sum_{i,j}^{occup} \int\int d\mathbf{r} d\mathbf{r}' \phi^{*}_{i}(\mathbf{r})\phi_j(\mathbf{r}) v(\mathbf{r},\mathbf{r}') \phi^{*}_{j}(\mathbf{r}')\phi_i(\mathbf{r}') \tag{1.79}$$

and,

$$E^{RPA}_{c} = \frac{1}{2\pi}\int_0^{\infty} d\omega\, \mathrm{Tr}\big[ln(1-\chi^{0}(i\omega)v + \chi^{0}(i\omega)v\big] \tag{1.80}$$

Although, RPA is exact for exchange energy, it is not exact for correlation energy. However, RPA usually has an accurate long-range electron correlation which is important for capturing long-range dispersion interactions [66, 67]. RPA predicts correct dissociation of diatomic molecules and gives a better description of interaction energies of molecular dimers in the S22 data set [59], reaction barrier heights and adsorption of atoms and molecules on different metallic surfaces [67]. However, RPA underestimates the atomization energies of small molecules referred to as "the RPA atomization energy puzzle" [66]. RPA can be further improved (beyond RPA methods) through fully nonlocal corrections [68] often through the introduction of exchange-correlation kernels.

1.7 Machine Learning in DFT

Machine Learning (ML) is a subfield of computer science, with the ambitious goal of studying and designing intelligent systems. In recent years, we have witnessed rapid progress

in ML—these systems are now reaching human-level and even superhuman-level performance for a range of tasks such as speech recognition, image interpretation, machine translation (Google translate), and gameplay (*DeepBlue* and *AlphaZero* for Chess, and *AlphaGo* and AlphaZero for Go). There is also a general belief that ML is poised to radically transform many components of our society and economy. This tremendous progress is leading to a radical shift in ML research, from a mainly academic endeavor to a much broader field with increasing industrial and governmental investments [69].

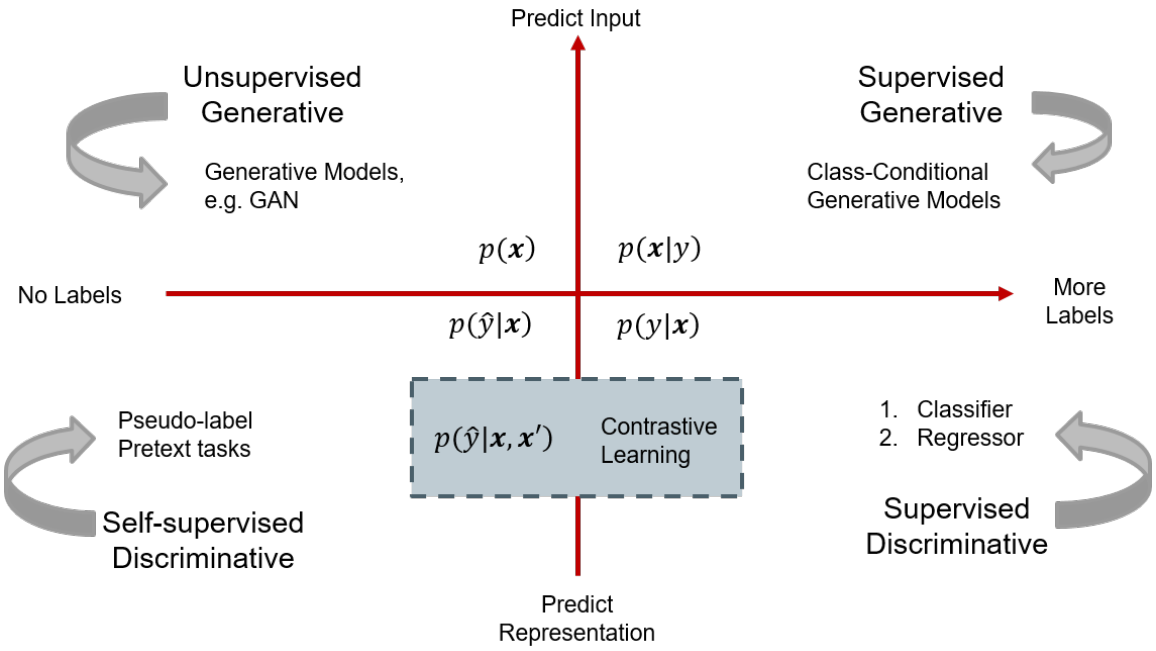

Figure 1.2: Generative-Discriminative and Supervised-Unsupervised spectrum. Contrastive learning belongs to the group of discriminative models that predict a pseudo-label of similarity or dissimilarity given a pair of inputs. [70]

Given many samples of random vectors $\mathbf{y}$ and $\mathbf{x}$, unsupervised learning attempts to approximate the joint probability distribution of $\mathbf{y}$ and $\mathbf{x}$ without making any assumptions about causality. Classification involves estimating the probability distribution for $\mathbf{y}$ given $\mathbf{x}$ (see Figure 1.2). The opposite operation, estimating the probability distribution of $\mathbf{x}$ given $\mathbf{y}$ is often called prediction (see Figure 1.2) when $\mathbf{y}$ causes $\mathbf{x}$ by being earlier data in a time sequence; in other cases where $\mathbf{y}$ causes $\mathbf{x}$, for example via a generative model, this operation is sometimes known as probability density estimation [71]. Generative approaches learn representations by modelling the data distribution $p(\mathbf{x})$, for example: all the pixels in an image. It is based on the assumption that a good model $p(\mathbf{x})$ that can generate realistic data samples, must also in turn capture the underlying structure related to the explanatory variables y. Evaluating the conditional distribution $p(y|\mathbf{x})$ for some discriminative tasks on

variable y can then be obtained by using Bayes' rule [70].

The application of machine learning techniques spread to the creation of exchange and correlation potential and energy functionals. The first application emerged from the work of Tozer et al. [72], where they devised a one-layer feed-forward multiperceptron neural network to map the electronic density $n(\mathbf{r})$ to the exchange and correlation potential $v_{xc}(\mathbf{r})$ at the same points. Technically, this exchange and correlation functional belongs to the family of local-density approximations. Tozer et al. [72] trained the neural network on two different datasets, first on the data of a single water molecule and afterwards on several molecules (namely, Ne, HF, N_2, H_2O, and H_2). Using 3768 data points calculated with a regular molecular numerical integration scheme [73], the method achieved an accuracy of 2–3% in the exchange and correlation energy of the water molecule. When applied in a self-consistent Kohn–Sham calculation, the potential lead to eigenvalues and optimized geometries congruent with the local density approximation. On the other hand, for the set of several molecules, Tozer et al. [72] obtained an error of 7.6% using 1279 points. The points were obtained in the same manner as before but were constrained to avoid successive points with similar densities. This potential generated geometries close to the local density approximation and good eigenvalues for molecules sufficiently represented in the training set. Meanwhile, and as expected, the neural network potential failed for molecules not sufficiently represented in the training like LiH and Li_2.

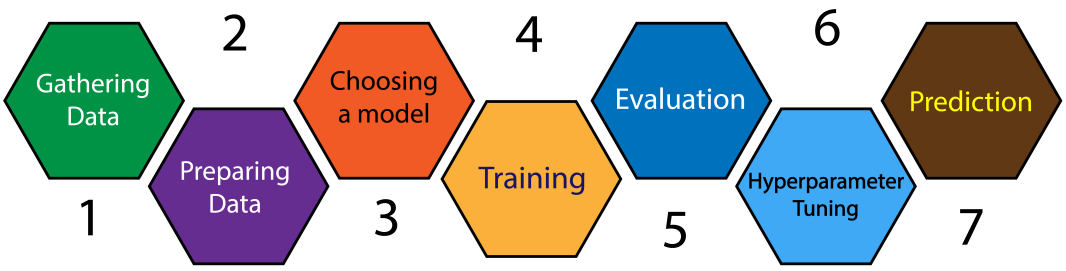

Figure 1.3: The General Steps of Machine Learning in Material Science

In 2012, Snyder et al. [74] tackled the problem of noninteracting spinless fermions confined to a 1D box. They employed KRR to construct a machine learning approximation for the kinetic energy functional of the density. This is the idea behind orbital-free DFT and an attempt to bypass the need to solve a Shrödinger-like equation. The kinetic energy and density pairs of up to four electrons were obtained using Numerov's method for several external potentials. These potentials were created using a linear combination of three Gaussian dips with random depths, widths, and centers. Furthermore, 1000 densities

were taken for the test set while M were taken for the training set. For $M = 200$ chemical accuracy was achieved, as no error surpassed 1 kcal/mol.

Recently, the DeepMind 21 (DM21) functional proposed by Kirkpatrick et al. [75] stands on the fully nonlocal fourth rung [76], which employs some exact exchange. The fractional constraints are translated into theoretical data that a machine can learn. DM21 is intriguing as separated atoms are properly neutral in DM21. This functional can describe all steps in a chemical reaction from reactants to transition state to products. DM21 is remarkably accurate for main-group chemistry, including some strongly-correlated systems like stretched H_2, which it describes correctly without the symmetry breaking that occurs on rungs one to three [76]. Because of its fractional constraints, DM21 might also describe strongly-correlated transition-metal oxides, although that remains to be seen.

On the other hand, Dick and Fernandez-Serra [77] took a different route and applied differentiable programming. Using an end-to-end differentiable implementation of the Kohn-Sham self-consistent field equations, they obtained an accurate neural network-based exchange and correlation (XC) functional of the electronic density. The machine was instructed to learn a SCAN-like density functional, with some of SCAN's exact constraints, from data. It is also possible to plot the enhancement factor over local exchange, and to show that SCAN is reasonably close to this optimal fit to the data. Differentiable programming allows the machine to construct not only the functional but also its first functional derivative, and thus to fit self-consistent DFT energies to accurate reference values [77]. Prof. John P. Perdew pointed out these issues in a private communication recently,

> *"I would however like to stress that machine learning, like other forms of data fitting, is primarily interpolative, while the satisfaction of exact constraints is primarily predictive. Perhaps the best functionals of the future will make the best of both worlds. I think that is what the Deepmind and Fernandez-Serra works point toward."*

CHAPTER 2

DENSITY FUNCTIONALS COMBINED WITH VAN DER WAALS CORRECTIONS FOR GRAPHENE ADSORBED ON LAYERED MATERIALS

2.1 Introduction

The concept of van der Waals (vdW) force stems from the famous work of Dutch physicist Johannes Diderik van der Waals on the equation of state for gases and liquids over 140 years ago [78, 79]. Intermolecular interactions are responsible for the behaviors of real gases and phase transitions between gas and liquid states. Although the vdW interaction loosely or generally refers to all intermolecular interactions [80], it more specifically refers to the London dispersion interaction [81], a universal interaction between any pair of objects made of electrons. The modern view of vdW interactions is that they are due to the correlation between fluctuations in electron densities [**l5**], or more specifically, the physics of collective plasmon oscillations [82]. Fluctuations are inherent in the electronic ground state [**l7**]. Correlated fluctuations are also responsible for another important interaction, namely the Casimir interaction, which works in a regime where the interacting objects are usually of macroscopic dimension and the separations between the interacting objects are on the sub-micron or micron scale, with retardation effects due to the non-zero propagation time of electromagnetic fields [**l5**, **l7**]. The vdW interactions take an electrostatic limit without retardation effects, and work in a regime where the separations between

interacting objects are on the sub-micron scale (usually $\mathring{A}$ to a few nm) [**l5**, **l7**]. Although the London-dispersion-based vdW interaction can be negligible in most solid systems, it has non-negligible effects in alkali metals [83], and plays a significant role in molecular complexes, molecular crystals, layered materials [**l9**], surface adsorptions, and many biosystems [84]. Casimir effects, which are found to be important in structured materials and composites with complex geometries, have also been drawing increased attention and studies in micromechanics [85], photonics [86], and complex biosystems [87, 88, 89].

Since an atomically-thin graphene film was first exfoliated and studied [90], 2-dimensional (2D) layered materials have become a very attractive and rapidly growing area in material science and technology, due to their many amazing properties and promising applications in the fields of energy recovery and storage [91, 92, 93, 94], catalysts [95, 96], next-generation electronics [97, 98] and optoelectronics [99]. 2D layered materials can be engineered by intercalations [95, 96, 97] and/or layer-on-layer assemblies in the so called vdW heterostructures [100, 91], giving them many opportunities to achieve various novel properties. Moreover, their electronic and optical properties are tunable via shapes or configurations [101, 102, 103], strains [104], and bending [105], etc. Recently, transition metal dichalcogenide (TMD) monolayers and multilayers, such as MoS_2, WS_2, $MoSe_2$, and WSe_2, have been emerging as important and widely studied 2D materials because of their many technologically favorable and novel properties [106], such as direct band gap (for monolayers), spin-orbital coupling splitting [107], and k-valley index [106], etc., making them promising materials for low cost, small sized, flexible, highly-efficient electronics and optoelectronics. When TMD monolayers or multilayers are stacked (or assembled) via vdW interactions on a graphene layer, which is conductive with zero gap and high carrier mobility, improved performances are usually achieved, such as reduced thermal transport [108], highly- efficient photon absorption and electron-hole pair creation [109, 110], making them suitable for high performance field-effect transistors [**37**, 111], transducers [112], and photodetectors [113], etc.

The dominant interactions between layers of 2D layered materials are long-range vdW interactions, which are many-body and nonlocal. The exact treatment for vdW interactions can be expressed in the adiabatic connection fluctuation-dissipation (ACFD) theorem [110, 114], in which the interacting many-body system is assumed to be adiabatically connected to a reference noninteracting mean-field system. However, the important interacting and non-interacting response functions in the ACFD expression have to be approximated in practice. The simplification of the interacting response function leads to the random phase approximations (RPA) [115, 116]. The ACFD or RPA based methods are not computationally efficient for large systems with over several hundreds of atoms per super-cell, and therefore many more efficient approaches [117, 118, 119] have been

developed within the framework of density functional theory (DFT) [120]. These include the vdW-DF family [117, 118, 119, 121], VV10 (Vydrov and Van Voorhis) [122] and rVV10 [123], DFT+D series [124, 125], and Tkatchenko-Scheffler (TS) methods [126, 127, 128]. The Kohn-Sham (KS) DFT methods [120] are most widely used in calculations of electronic structure of various materials and predictions of properties of materials, due to their high efficiency and reasonable yet improvable accuracy. Conventional DFTs approximate the exchange-correlation interactions locally [35] or semilocally. Although the advanced semilocal functional SCAN (Strongly Constrained and Appropriately Normed) [40] can capture intermediate correlation interactions, the long-range correlation effects, which largely account for the vdW exfoliation energies of the layered materials [**l9**], are still not included in conventional DFT methods. The DFT-vdW combined methods are perhaps the most efficient way to include vdW interactions. However, most of the vdW methods are based on a pairwise interacting picture. They are developed via the asymptotic formula $-C_6/D^6 - C_8/D^8$. Many methods have been devoted to the leading-order term, while the higher-order contribution (e.g., C_8 term) is based on the leading-order coefficient C_6, as in DFT+D3 [64], where the important dynamic dielectric screening effects are not explicitly included. Ruiz et al. [129] developed a DFT+vdWsur method, in which the coefficient C_6 is renormalized with the Lifshitz-Zaremba-Kohn (LZK) theory [130] for atom-surface interactions and can include the Coulomb screening within the surface implicitly. Although the high order terms are not included in DFT+vdWsur, it still gives very accurate results for several hybrid inorganic-organic systems.

The asymptotics of the vdW energy of two interacting entities of matter depend on their geometries and electronic properties. The vdW energy of two parallel slabs of infinite thickness, separated by distance D , is inversely proportional to D^2 [131] in the large D limit. Dobson et al. [132] found that the asymptotic vdW energy between a thin metallic layer and a parallel-placed -conjugated thin layer goes as $\sim D^{-3} log(D/D_0)$, where D_0 is a constant related to Fermi energy. However, from a random phase approximation analysis, Lebègue et al. [133] showed that the asymptotic behavior of the interlayer dispersion interaction of graphite varies as D^{-3} . The leading term in the vdW interaction between a particle, i. e., an atom, and a surface is $-C_3/D^3$ [**71**]. A more general form including quadrupole-surface contribution is [133, 134].

$$E_{vdW} = -\frac{C_3}{D^3} - \frac{C_4}{D^4} - \frac{C_5}{D^5} \tag{2.1}$$

where $C_4 = C_4^d + C_4^{nl}$, C_4^d is the diffuse term and C_4^{nl} is the nonlocal term. The term C_5 consists of three contributions, $C_5 = C_5^q + C_5^{nl} + C_5^d$, where C_5^q represents the quadrupole term. Zaremba and Kohn (ZK) introduce a quantity Z_0 [129], representing the position of a reference plane for the vdW interaction between an atom and a solid surface from a

second-order perturbation analysis, so that Z_0 will zero out the C_4 term in Eq. 2.1, but not the C_5 term, leading to $E_{vdw} = -C_3/(D-Z_0)^3 - C_5/(D-Z_0)^5$. For a jellium substrate, D is the distance from the background edge of the jellium surface to the atom [134].

Based on the ZK theory, we have developed a vdW model [135, 136], denoted as vdW-dZK (van der Waals damped Zaremba-Kohn), for incorporating the important dynamic dielectric screening effects from the solid surfaces in evaluations of the C_5 term. The developed vdW model has been combined with DFT PBE (Perdew-Burke-Ernzerhof) calculations and successfully applied to graphene and thiophene adsorbed on various metal surfaces [136, 137]. It predicts accurate binding energies and equilibrium distances between the adsorbates and metal surfaces, and the binding energy curves obtained for graphene on metals that are also very close to those from RPA calculations around the equilibrium distances. One key point in our vdW-dZK model is that we treat the graphene layer and thiophene molecule as a collection of individual atoms, each with its renormalized dynamic multipole polarizability. That means that we ascribe a fraction of the polarizability of the molecule to each of its constituent atoms, with the fraction being the same as it would be for a collection of separated atoms.

In this work, with some modifications, we extend the developed vdW model with DFT calculations to graphene adsorbed on monolayer and multilayer layered materials, such as MoS_2,WS_2,$MoSe_2$,WSe_2, graphite and h-BN. Since most of the studied layered materials are semiconducting, the background edge of jellium surfaces is hardly applicable and the C_4 term for semiconducting materials may not be small. We use Eq. 2.1 in this work. In addition, the multilayer TMD is treated as a slab of a uniform dielectric continuum. The effects of multiple images of dipole and quadrupole, formed in space by the two surfaces of the slab, will also be considered.

2.2 Theoretical Model

We set a dipole moment p, $p = ql$, in front of a dielectric slab with a thickness c. The distance from the center of the dipole to the surface is denoted as a. The direction of the dipole is perpendicular to the surface and pointed away from the surface. The origin of the coordinate system is at the center of the dipole, as shown in Fig. 2.1. The front and back surfaces of the slab are denoted as S_a and S_b, respectively. b is the distance from the origin to the back surface and $b = a + c$. As analyzed by Sometani [138], each of the positive and negative charges in the dipole will form four sets, each an infinite number of image charges, denoted as $q_{aan}, q_{abn}, q_{bbn}$, and q_{ba}, somewhere on the Z axis, where n in the subscript labels the nth image and can be any integer number from 1 to infinity. The first two sets, $q_{aan} and q_{abn}$, are equivalent to two sets of induced surface charges distributed on the front

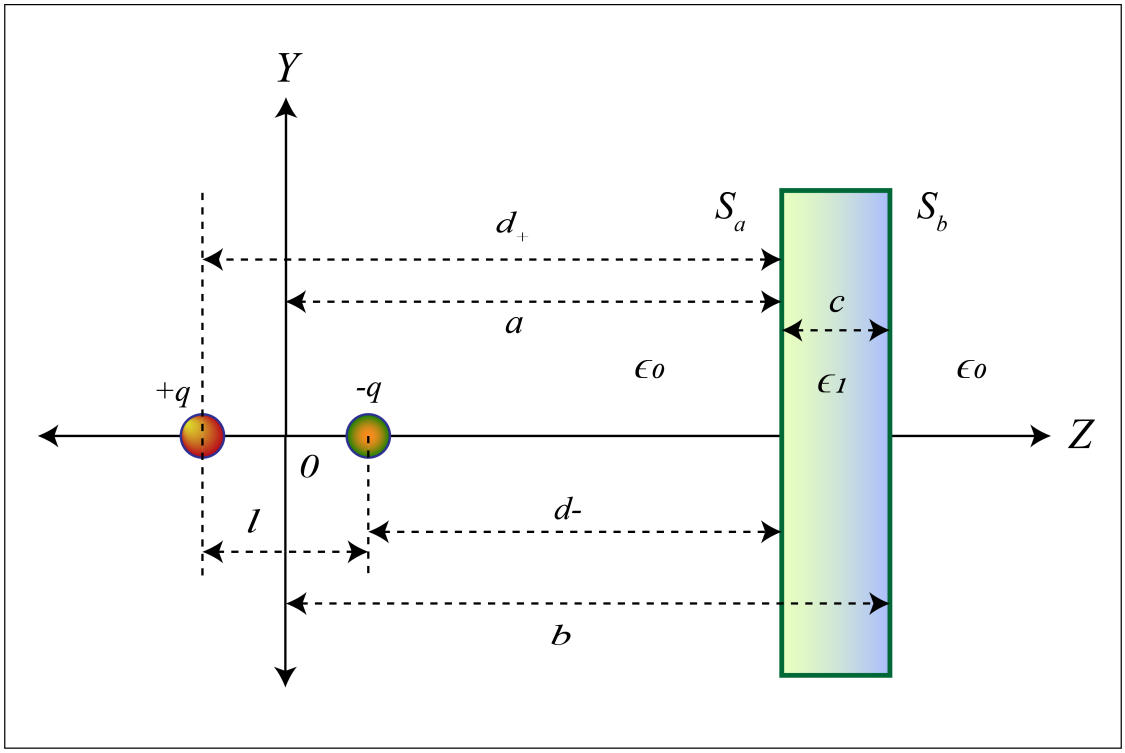

Figure 2.1: Schematic diagram of an interacting system of a dipole and a dielectric slab.

surface S_a, and are associated with fulfillment of electrostatic boundary conditions on the front surface S_a and on the back surface S_b, respectively. Similarly, the other two sets, q_{bbn} and q_{ban}, are equivalent to another two sets of induced surface charges distributed on the back surface S_b, and are associated with fulfillment of electrostatic boundary conditions on the back surface S_b and on the front surface S_a, respectively.

The total vdW interaction energy, $E_{vdW}^{total,dipole}$, between the dipole moment p and the dielectric slab can be calculated as a sum of two parts, where the first part corresponds to the total electrostatic potential energy of the positive charge in the dipole moment p interacting with all the images in the four sets associated with not only the positive charge but also the negative charge in the dipole moment p, and similarly, the second part corresponds to that of the negative charge in the dipole moment p interacting with all the images in the four sets. For a given charge q, located at a distance d in front of the surface S_a, the electrostatic potentials, not potential energies, produced by its images in the four sets, on a point Z on the Z axis are [138]

$$V_{aan} = \frac{q_{aan}}{|Z-a|+(2n-2)c+d} \tag{2.2}$$

$$V_{abn} = \frac{q_{abn}}{|Z-b|+(2n-1)c+d} \tag{2.3}$$

$$V_{bbn} = \frac{q_{bbn}}{|Z-b|+(2n-1)c+d} \tag{2.4}$$

$$V_{ban} = \frac{q_{ban}}{|Z-a|+2nc+d} \tag{2.5}$$

where $q_{aan} = \beta^{2n-1}(-q_0)$, $q_{abn} = \beta^2 n(-q_0)$, $q_{bbn} = \beta^{2n-1} q_0$, $q_{ban} = \beta^{2n} q_0$, $q_0 = q$, $\beta = \frac{\varepsilon_1 - 1}{\varepsilon_1 + 1}$, ε_1 is the relative dielectric constant of the slab. $V_{aan}, V_{abn}, V_{bbn}$, and V_{ban} are the four sets of electric potential produced by the image charges of $q_{aan}, q_{abn}, q_{bbn}$, and q_{ban}, respectively. Atomic units ($e^2 = m_e = \hbar = 1$) are used through the paper unless otherwise noted.

The location of the positive charge is $Z = -l/2$, while that of the negative one is $Z = l/2$. With these and Eqs. 2.2, the total vdW energy between the dipole and the slab is written as,

$$\begin{aligned} E_{vdW}^{total,dipole} = \sum_n \{ & q\left[(V_{aan}+V_{abn}+V_{ban}+V_{bbn})|_{Z=-l/2,d=d_+,+q}\right. \\ & \left.+(V_{aan}+V_{abn}+V_{ban}+V_{bbn})|_{Z=-l/2,d=d_-,-q}\right. \\ & -q\left[(V_{aan}+V_{abn}+V_{ban}+V_{bbn})|_{Z=l/2,d=d_+,+q}\right. \\ & \left.+(V_{aan}+V_{abn}+V_{ban}+V_{bbn})|_{Z=l/2,d=d_-,-q}\right. \end{aligned} \tag{2.6}$$

Since l is much smaller than a, b, and c, we can expand Eq. 2.6 to the second order of l (with the full derivation to be found in Appendix A),

$$\begin{aligned} E_{vdw}^{total,dipole} = & -2q^2\beta^{(2n-1)}\left(l^2/\Xi^3\right) - 2q^2\beta^{2n}\left(l^2/\xi^3\right) \\ & +2q^2\beta^{2n}\left(l^2/\xi^3\right) - 2q^2\beta^{(2n-1)}\left(l^2/\xi^3\right) \end{aligned} \tag{2.7}$$

where, $\Xi = 2a+(2n-2)c$ and $\xi = a+b+(2n-1)c$. And, $b = a+c$ gives us $\xi = 2a+2nc$

In Eq. 2.7, the second and the third terms cancel each other. This means that the contributions from V_{abn} and V_{ban} are cancelled. When n=1, the first term in Eq. **??** is $\frac{-2^2 l^2}{(2a)^3}$, which is just the vdW interaction of a dipole with a slab of infinite thickness and is equivalent to the first term in Eq. 2.1 in ref. [136] or in Eq.(1.2) in ref. [139]. This means that the factor $\frac{2^2 l^2}{2^3}$ is equivalent to $\frac{1}{4\pi}\int_0^\infty du\ \alpha_1(iu)\ \frac{\varepsilon_1(iu)-1}{\varepsilon_1(iu)-1}$. Note that the former factor is derived from a classic electrostatic approach, while the later expression is derived from second-order perturbation theory in quantum mechanics [130]. With this equivalence, Eq. 2.7 can be written as,

$$E_{vdw}^{total,dipole} = \sum_n \left(\frac{-1}{(a+(n-1)c)^3} + \frac{1}{(a+nc)^3}\right) C_{3,n} \tag{2.8}$$

where the coefficient, $C_{3,n}$, of dipole-surface interaction for the n-th image is expressed as,

$$C_{3,n} = \frac{1}{4\pi}\int_0^\infty du\ \alpha_1(iu)\ \frac{\varepsilon_1(iu)-1}{\varepsilon_1(iu)-1}\beta^{2n-2} \tag{2.9}$$

where iu is the imaginary frequency and $\beta = \frac{\varepsilon_1(iu)-1}{\varepsilon_1(iu)-1}$. The mean dynamic multipole polarizability of the adsorbate particles, i.e., carbon atoms in the graphene layer, is $\alpha_t(iu) = \alpha_t(0)\frac{\omega_t^2}{(\omega_t^2+u^2)}$ with $\omega_t = \omega_p\sqrt{t/(2t+1)}, \omega_p = \sqrt{4\pi m}$ and $\alpha_t(0) = [\alpha_1(0)]^{(2t+1)/3}$. (Note that $t = 1$ for dipole and $t = 2$ for quadrupole.) The isotropic dielectric function $\varepsilon_1(iu)$ of the slab will be given in Eq. (17). The average electron density of graphene is $m = 0.13$ atomic unit. The static dipole polarizability per atom of graphene $\alpha_1(0)$ is taken to be 9.945 atomic unit [140].

A similar analysis applies to the quadrupole term C_5^q, since the electric potentials produced by the images of each charge in a quadrupole are also given by Eqs. 2.2-2.5. Moreover, since the diffusion terms C_4^d and C_5^d are due to smoothly-varying electron density profiles of surfaces [134], and are related to dipoles and quadrupoles, respectively, the image contributions for C_4^d and C_5^d terms can also be accounted for by the images of dipoles and quadrupoles, and are also similar to Eq. 2.8. The nonlocal terms C_4^{nl} and C_5^{nl} [134] are derived from the hydrodynamic model at electrostatic limits, and are actually dipolar contributions contributing to the terms of $\sim 1/D^4$ and $\sim 1/D^5$. The nonlocal quadrupole contribution is more complicated and not considered here. The image contributions for C_4^{nl} and C_5^{nl} terms can also be included by the images of each charge in the dipole. Eventually, we get the total vdW interactions between a particle and a surface as,

$$\boxed{E_{vdw}^{total} = \sum_n \sum_{i=3}^{5} \left(\left(\frac{-1}{(a+(n-1)c)^i} + \frac{1}{(a+nc)^i} \right) C_{i,n} \right) f_{damp,n}} \tag{2.10}$$

where $C_{3,n}, C_{4,n}$ and $C_{5,n}$ (for $i = 3, 4$, and 5, respectively) are coefficients relating to the nth image. Letting $n = 1$ and $c \to \infty$, Eq. 2.10 reduces to Eq. 2.1 with a leading a^{-3} decay at large separation a. But, for finite thickness c and $a \to \infty$, the quantity in the large brackets in Eq. 2.10 tends to $-3c/a^4$. Note that $C_{4,n} = C_{4,n}^d + C_{4,n}^{nl}$ and $C_{5,n} = C_{5,n}^q + C_{5,n}^d + C_{5,n}^{nl}$. The damping function $f_{damp,n}$ [135] is included to remove spurious divergence in Eq. 2.10 when a is very small. $f_{damp,n}$ is given by,

$$f_{damp,n} = \frac{x^5}{\sqrt{1+gx^2+hx^4+x^{10}}} \tag{2.11}$$

where $x = ((a+(n-1)c))/\bar{b}$, $g = 2\bar{b}^2 C_{3,n}/C_{5,n}$ and $h = 10\bar{b}^4 C_{3,n}^2/C_{5,n}^2$. $\bar{b}$ is the cutoff parameter for the damping function. The coefficient $C_{5,n}^q$ is calculated as

$$C_{5,n}^q = \frac{1}{4\pi}\int_0^\infty du\ \alpha_2(iu)\ \frac{\varepsilon_1(iu)-1}{\varepsilon_1(iu)-1}\beta^{2n-2} \tag{2.12}$$

The nonlocal parts are given as,

$$C_{4,n}^{nl} = -\frac{3}{8\pi}\int_0^\infty du\ \alpha_1(iu)\ \frac{\varepsilon_1(iu)(\varepsilon_1(iu)-1)\xi}{(\varepsilon_1(iu)-1)^2}\beta^{2n-2} \tag{2.13}$$

$$C_{5,n}^{nl} = \frac{1}{4\pi}\int_0^\infty du\ \alpha_1(iu)\ \frac{\varepsilon_1(iu)(\varepsilon_1(iu)-1)^2\xi^2}{(\varepsilon_1(iu)-1)^3}\beta^{2n-2} \tag{2.14}$$

where $\xi = \delta/\sqrt{\bar{\omega_p}^2+u^2}$, $\delta = 3k_F^2/5, \bar{\omega_p} = \sqrt{4\pi\bar{n}}$, $\bar{n}$ valence electron density of the slab, and $k_F = (3\pi^2\bar{n})^{1/3}$ the average Fermi wavevector. The diffusion parts are evaluated as,

$$C_{4,n}^{d} = \frac{\alpha_1(0)\omega_1^2\omega_s^2}{2}\int_0^\infty du\ g_0\ f(u) d_\perp(iu)\beta^{2n-2} \tag{2.15}$$

$$C_{5,n}^{d} = \frac{\alpha_1(0)\omega_1^2\omega_s^4}{2(\omega_1^2+\omega_s^2)}\int_0^\infty du\ g_0\ f(u) d_\perp^2(iu)\beta^{2n-2} \tag{2.16}$$

where,

$$f(u) = \frac{[u^2+(1+g_0)\omega_s^2]}{[(u^2+\omega_1^2)(u^2+\omega_s^2)^2]} \tag{2.17}$$

$$g_0 = (1+u^2/\omega_s^2)\left(\frac{\varepsilon_1(iu)-1}{\varepsilon_1(iu)-1}\right) \tag{2.18}$$

$$d_\perp(iu) = d_\perp(0)\left[1-\lambda\ tan^{-1}(1/\lambda)\right] \tag{2.19}$$

$$\lambda^2 = \frac{[u^2+\omega_s^2(1-g_0)]}{2\omega_s^2 g_0} \tag{2.20}$$

and $\omega_s^2 = \omega_p^2/2$. ω_1 and ω_p are defined after Eq. 2.9. The static centroid of the induced charge density on the surface is parametrized as,

$$d_\perp(0) = 0.02r_s^2 - 0.27r_s + 2.06 \tag{2.21}$$

with $r_s = [3/(4\pi\bar{n})]^{1/3}$

The dynamic dielectric function, $\varepsilon_1(iu)$, of the semiconducting slab is modeled by the modified Penn model as [141],

$$\begin{aligned}\varepsilon(iu) = 1 + \frac{\omega_p^2}{u^2}\left[\frac{(1-\Delta^2)\,y}{P} - \frac{\omega_g^2-(\omega_g^2+u^2)\,\Delta^2}{2u\sqrt{(\omega_g^2+u^2)}}log\frac{I_+}{I_-}\right] \\ +\frac{2\omega_p^2\Delta}{u^2}\left\{\frac{\omega_g}{u}\left[tan^{-1}\left(\frac{\omega_g P}{u}\right) - tan^{-1}\left(\frac{\omega_g}{u}\right)\right] + \frac{1}{P} - 1\right\}\end{aligned} \tag{2.22}$$

where $I_\pm = \left[\left(1+y^2\right)\left(1+u^2/\omega_g^2\right)\right]^{1/2}, y = 1/\Delta$, and $P = \sqrt{1+y^2}$ ω_g is the effective energy gap, which is defined below, and $\Delta = \omega_g/(4\varepsilon_F)$ with $\varepsilon_F = (3\pi^2 n_{nanotube}^{2/3})/2$, with $n_{nanotube}$ defined in the second paragraph after Eq. 3.15. This model dielectric function has been used to study physical adsorption [142, 143, 144]. The effective energy gap ω_g is related to the static dielectric function $\varepsilon(0)$ via the Penn model [145],

$$\varepsilon(0) = 1 + \frac{2\omega_p^2}{3\omega_g^2}[(1+\Delta^2)^{1/2} - \Delta] \tag{2.23}$$

Since the accurate $\varepsilon(0)$ from a high-level *ab initio* calculation contains important hybridization, inhomogeneity, and exciton effects, which can be carried over to our model via the Penn model, the model dielectric function is expected to work well for semiconductors [144]. The modeled dielectric functions for *Si* and *GaAs* agree very well with the corresponding GW and GW+BSE results [146], and are computationally very efficient. The effective energy gap ω_g is determined by reproducing the $\varepsilon_1(0)$ value of Eq. 2.23. This energy gap is different from the usual band gap. For example, for the monolayer WS_2, the isotropic static dielectric function is $\varepsilon_1(0) = (24.4 + 2.9)/3 = 3.9$ atomic unit [147, 148]. By Eq. 2.23, this leads to the effective energy gap $\omega_g = 0.445$ atomic unit.

When graphene interacts with a multilayer TMD, it must interact primarily with the negative chalcogen ions, so we take the thickness of the slab to be the distance from the topmost to the bottommost chalcogen plane (making D of Eq. 2.1 equal to a of Fig.2.1. The monolayer WS_2 is treated as a slab, and the average valence electron density $\bar{n}$ of the slab is calculated as $n = (2+26)/((3.15/0.529)^2(3/2)(3.1/0.529) = 0.078$ atomic unit, where the intralayer lattice constant is 3.15 $\AA$ [149, 150] and the thickness of the WS_2 monolayer slab is taken to be 3.1 $\AA$ [149, 150], one quarter of the bulk WS_2 interlayer lattice constant c. The numbers of valence electrons are 2 for the W atom and 6 for the S atom, respectively. For bilayer WS_2, $n = 2(2+26)/((3.15/0.529)^2(3/2)(9.3/0.529) = 0.052$ atomic unit, because there are 2 W atoms and 4 S atoms in the cell and the thickness of the slab is also increased by half of the bulk WS_2 interlayer lattice constant c. The same scheme is used for evaluating the average valence electron density $\bar{n}$ of the 4-layer WS_2 slab and other layered materials. We have used experimental geometries for the substrates, although SCAN geometries might suffice.

The modified vdW model, Eq. 2.10, is combined with two DFT functionals, namely, SCAN and PBE, to evaluate the binding energies and equilibrium distances for graphene (G) parallelly adsorbed on the planar surfaces of several monolayer and multilayer layered materials. All DFT calculations are done in the Vienna Ab initio Software Package (VASP) [151] with projector augmented-wave (PAW) pseudopotentials [152, 153]. The energy cutoff is 580 eV. A 5×4 unit cell is used for the computational cells of G/MoS_2 and G/WS_2, a 4×3 unit cell for that of G/$MoSe_2$ and WSe_2, and a 1×1 unit cell for that of G/Graphite and G/h-BN_{bulk}.

The computational cells are so designed that the lattice mismatch between graphene and substrates is less than 3%. The thickness of the vacuum layer is about 25 $\AA$. The k-point mesh of $4\times4\times1$ is used for the cases of G/MoS_2 and G/WS_2, G/$MoSe_2$ and WSe_2, while a k-point mesh of $20\times20\times1$ is used for G/Graphite and G/h-BN. Bulk graphite and h-BN are modelled with a 6-layer slab.

All structures of graphene on layered materials are optimized with each DFT func-

tional and then the DFT curves of binding energy vs. distance $E(D)$ are calculated. The binding energy is defined as the total energy of the graphene-substrate structure after subtraction of the energy of the separated graphene layer and the energy of the separated substrate. D is the direct distance between the graphene layer and the first layer of atoms in the layered materials. The curves of vdW energy vs. distance are computed in Python via Eq. 2.10 and are combined with the DFT $E(D)$ curves to get the total binding energy curves, whose minima are equilibrium points, hence the determinations of the equilibrium binding energies and equilibrium distances. The flow chart of the computational process is shown in Fig. 2.2. Several Python scripts have been developed to perform the tasks. For comparison, the results with the SCAN+rVV10 [**l9**] and PBE+rVV10 [137, 154] calculations are also obtained.

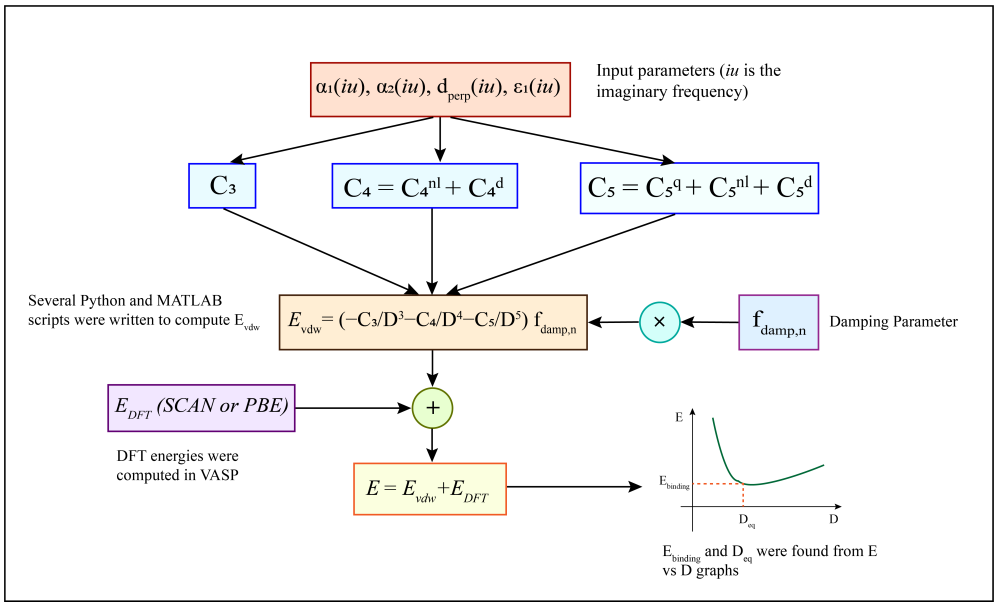

Figure 2.2: Flow chart of the computational process for the vdW interaction energies, DFT energies, and total energies.

2.3 Results and Discussions

The results for binding energies and distances at equilibrium from different methods are shown in Table I.

The RPA value for the binding energy of graphene on the surface of bulk graphite is -48.0 meV/C atom [132]. The RPA equilibrium distance for G/graphite is 3.34 $\AA$ [149, 150]. Reference data for the other four layered materials (MoS_2,WS_2,$MoSe_2$, and WSe_2) considered here are scarce. For G/MoS_2 monolayer, the binding energies are reported

as $-21 \sim -25$ meV/Carbon atom, and the equilibrium distances are about $3.3 \sim 3.6$ $\AA$ [155, 156, 157]. For G/WS_2 and G/WSe_2, ref. [158] reports ~ 53 meV/C-atom, ~ 3.4 $\AA$. However, ref. [159] reports $\sim 21 - 32$meV/C-atom, ~ 3.9 $\AA$ for G/WS_2.

It is generally thought that the RPA includes both nonlocal self-interaction corrections and the full range of van der Waals corrections (vdW) [160], although it may underestimate the strength of chemisorption [161]. The RPA overestimation of the strength of short-range correlation does not seem to matter for vdW bond predictions. The RPA value for the physisorption energy would still be reasonably reliable and could be used as a temporary reference for vdW interactions, in the absence of better references. In references [162, 163], RPA values are used as reference values for solid layered vdW systems. Moreover, much progress has been made to improve the RPA or go beyond RPA. Many of these refinements improve the description of short-range correlation, while maintaining good performance for dispersive interactions, see [164, 165, 166]. Those developments provide improved predictions for many solid properties, such as lattice constants, bulk moduli, and cohesive energies, etc.

To the authors' best knowledge, RPA or higher-level results are not available for many systems considered here, due to the relatively large (about 100-250 atoms in supercells) systems. The TS+rsSCS-MBD [168] method calculates the vdW dispersion energy with the help of a coupled-fluctuating-dipole model Hamiltonian, in which each atom is modelled as a quantum harmonic oscillator whose characteristic frequency and frequency-dependent (dynamic) polarizability are obtained by a series of renormalized schemes and by solving the so-called self-consistent many-body screening equations. However, due to numerical issues, the TS+rsSCS-MBD in the VASP code is not suitable for the systems considered here at this moment. Gould et al. [162] reported revised Many Body Dispersion (MBD) (MBD@rsSCS-ER and MBD@rsSCS/FI) methods, which give very similar results for energy curves for some transition metal dichalcogenite 2D materials in comparison with the RPA. The newly developed uMBD by Kim et al, [163] is based on previous MBD@rsSCS-ER and MBD@rsSCS/FI methods. The uMBD gives even better agreement of binding energy results for some 2D layered materials with RPA results, see Fig. 4 in [163]. Also note that the vdW-DF-cx [169] and SCAN+rVV10 give larger interlayer binding energies for MoS_2 and hBN than do the RPA and uMBD. The results of vdW-DF-cx are even larger than those from SCAN+rVV10. This is shown by our calculations with vdW-DF-cx in Quantum ESPRESSO [**l113**, 170] for G-MoS2-1L, where at equilibrium, the binding energy is -54.7meV/C and the equilibrium distance is 3.37 angstroms from vdW-DF-cx, while the results from SCAN+rVV10 are -42.7meV/C and 3.46 angstroms, respectively. Further extensive comparison tests of various vdW methods will be a direction for future research. The currently available RPA values can only serve as an approximate

Table 2.1: The binding energies E_b (meV/C atom) and equilibrium distances D_{eq} (Angstrom) between a graphene layer (G) and substrates of TMD layered materials calculated from different methods. 1L stands for monolayer TMD substrate, 2L for bilayer TMD substrate, and 4L for tetralayer TMD substrate. The reference RPA values for G/Graphite are E_b = -48.0 and $D_{eq} = 3.34$. The reference RPA values for G/h-BN_bulk are E_b = -41.5 and D_{eq} = 3.35. The reference RPA values for all other systems can be found in Ref. [167]. D is the distance between the plane of graphene nuclei and the first plane of substrate nuclei, where the planes are defined by averaging the z coordinates of the corresponding nuclei.

	SCAN		SCAN + rVV10		SCAN + dZK		PBE		PBE+rVV10		PBE+dZK	
	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}
G/hBN_bulk	-22.41	3.41	-55.51	3.29	-44.94	3.37	-1.23	4.47	-43.90	3.46	-39.24	3.59
G/Graphite	-19.40	3.47	-52.32	3.36	-46.82	3.40	-0.79	4.47	-42.25	3.54	-46.12	3.55
G/MoS2_1L	-14.63	3.56	-42.70	3.46	-29.11	3.51	-1.51	4.39	-38.17	3.56	-29.69	3.51
G/MoS2_2L	-14.75	3.55	-44.80	3.43	-32.79	3.50	1.51	4.39	-38.17	3.56	-29.69	3.51
G/MoS2_4L	-14.76	3.55	-45.00	3.43	-34.60	3.49	-1.51	4.37	-40.84	3.54	-40.02	3.43
G/WS2_1L	-14.27	3.56	-42.41	3.47	-27.89	3.54	-1.53	4.39	-28.50	4.29	-28.08	3.58
G/WS2_2L	-14.41	3.55	-44.53	3.44	-31.44	3.51	-1.54	4.38	-40.42	3.56	-35.66	3.46
G/WS2_4L	-14.45	3.55	-44.95	3.44	-33.83	3.49	-1.54	4.38	-40.88	3.57	-39.77	3.43
G/MoS2_1L	-13.75	3.66	-40.44	3.52	-31.81	3.61	-1.20	4.49	-35.51	3.51	-33.96	3.56
G/MoS2_2L	-13.91	3.64	-42.32	3.52	-31.92	3.60	-1.20	4.49	-38.18	3.63	-35.89	3.49
G/MoS2_4L	-13.92	3.64	-42.82	3.51	-32.41	3.58	-1.20	4.49	-38.26	3.64	-36.31	3.48
G/WSe2_1L	-13.38	3.68	-40.25	3.52	-26.18	3.65	-1.17	4.50	-35.57	3.51	-24.86	3.64
G/WSe2_2L	-13.45	3.67	-42.23	3.52	-29.54	3.61	-1.18	4.48	-38.02	3.63	-31.09	3.59
G/WSe2_4L	-13.47	3.67	-42.21	3.52	-31.30	3.60	-1.18	4.48	-38.44	3.62	-33.95	3.55

reference, with which our SCAN-vdW-dZK values are in reasonable agreement.

The PBE GGA, which captures no long-range and very little intermediate-range van der Waals interaction, yields an unphysically weak binding energy of 1 or 2 meV/C atom, and an unphysically long equilibrium distance of about 4.5 $\AA$. The SCAN meta-GGA, which captures intermediate-range van der Waals interaction (on the length scale of about 3 $\AA$), yields only about half of the binding energy, but a rather realistic equilibrium distance. Much better binding energies are achieved by adding the long-range van der Waals corrections rVV10 or vdW-dZK to PBE or SCAN. However, SCAN+rVV10 and (to a lesser extent PBE+rVV10) clearly overbind. Note that the overestimate in binding energies by SCAN+rVV10 has also been found for G/Graphite and for a thiophene molecule adsorbed on metal surfaces [137]. The most plausible binding energies and equilibrium bond lengths come from SCAN+vdW-dZK and PBE+vdW-dZK. Generally speaking, SCAN-vdW-dZK and PBE-vdW-dZK give approximately the same results, especially for G on monolayer substrates. This indicates the consistency of our vdW-dZK method. The binding energies for G/monolayer predicted by SCAN-vdW-dZK and PBE-vdW-dZK are about $\sim$24-29 meV/C and are close to the results from other methods in the literature [171, 155, 156].

SCAN gives very similar results for each system when the substrate varies from monolayer to 4-layer. For example, for G/MoS_2, the binding energy E_b changes slightly from -14.6 to -14.8 meV/C, and the equilibrium distance D_{eq} is almost unchanged from 3.56 to 3.55 $\AA$. This indicates that, although SCAN improves a lot over PBE by capturing intermediate correlation interactions, it is still insufficient for the long-range correlation effects in the exfoliation energies of the layered materials, as found in ref. [**l9**]. When coupled with vdW correction rVV10, the change in binding energy is about 2 3 meV/C from monolayer through 4-layer substrates, indicating that more long-range correlation effects are included in SCAN+rVV10. From G/monolayer through G/4-layer, the change in binding energies is $\sim$1-5 meV/C from SCAN-vdW-dZK, $\sim$2-3 meV/C from SCAN+rVV10, $\sim$1-3 meV/C from PBE+rVV10, and $\sim$3-10meV/C from PBE-vdW-dZK. All approximations predict gradually stronger vdW interactions when the substrates change from monolayer through 4-layer.

The SCAN-vdW-dZK cutoff parameter $\bar{b}$ for the damping function is 4.50 Bohr (determined by best fit to the RPA binding energy value of G/Graphite bulk) and is relatively large, due to the capture of intermediate correlations by SCAN. The b of 3.30 Bohr for PBE-vdW-dZK is the same as that in our previous work [136], further confirming the consistency of our PBE-vdW-dZK method. The effects of the image and the back surface of the slab can be understood via Eq. 2.8 and Eq. 2.10. For example, for G/MoS_2, $C_{3,2}/C_{3,1}$ is about 0.3, the thickness of the slab c for monolayer MoS_2 is about the same as the equi-

librium a , say, $\sim 3.3\,\AA$, so the leading C_3-term contribution from the second image $(n=2)$ is about 5% of that from the first image $(n=1)$. The contribution of V_{bbn}, which is the second term on the right hand of Eq. 2.8 and is due to the back surface of the slab, can be estimated as $\sim 1/8$, about 12% of the first term on the right hand of Eq. 2.8, at $n=1$. However, for high order images and/or for thicker slab, like 2-layer and 4-layer substrates, the V_{bbn} contributions are quickly reduced and can be negligible. By a similar analysis, the contributions of C_3, C_4, and C_5 terms are about 50%, 40%, and 10%, respectively. The damping factor $f_{damp,1}$ at equilibrium is about 0.7 for the PBE-vdW-dZK and 0.3 for the SCAN-vdW-dZK, mainly due to the significant change of the cutoff parameter $\bar{b}$ from 3.3 to 4.5 Bohrs.

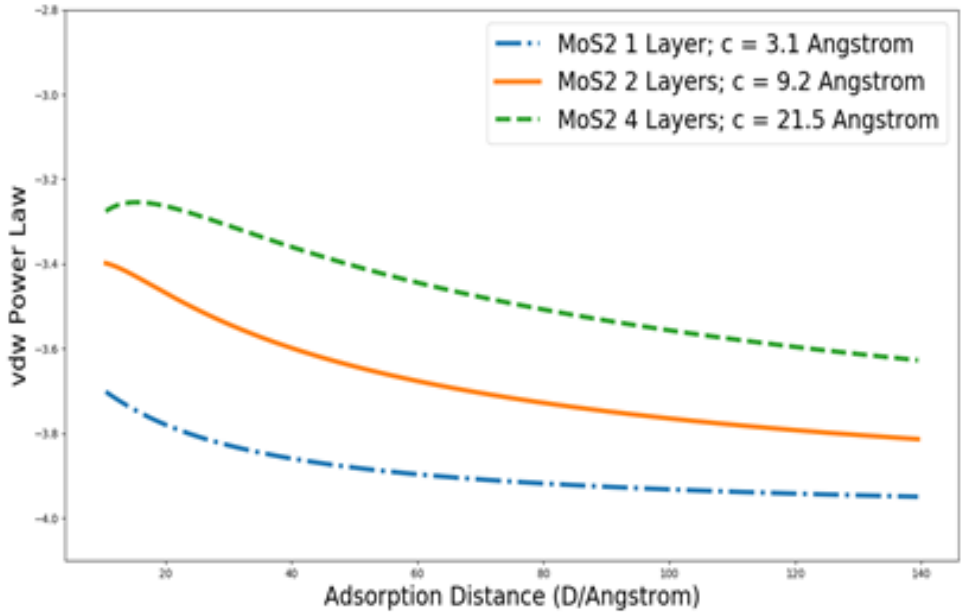

Figure 2.3: vdW interaction energy power law exponents from the vdW-dZK model for graphene adsorbed on MoS_2 multilayer as a function of adsorption distance $D=a$.

Ambrosetti et al. [172] analyzed vdW systems of atoms or small molecules and 2D thin metallic or finite-gap substrates with RPA-like and CDH (coupled dipolar Hamiltonian) treatments for substrate responses, and revealed nontrivial variations of power-laws of the vdW interactions in these systems. They found that these power laws substantially deviate from standard pairwise predictions, and result in ultra-long-ranged dispersion forces. The power law exponent $P(D)$ is defined as $P(D) = d(log|E_{vdw}|)/dlogD$, and is a slowly-decaying function of separation $D=a$. This effect arises due to marked nonlocalities of the substrate response.

We also calculated the vdW interaction energy power laws from our vdW-dZK model for graphene adsorbed on MoS2 multilayer. Fig. 2.3 shows the vdW power law exponent from our vdW-dZK model as a function of D. As can be seen, the vdW-dZK exponent is close to -3 at small D and tends slowly to -4 at large D. (See our discussion after Eq. 2.10.) These slowly varying features are also similar to those obtained from RPA and RG (renormalization group) approaches by Ambrosetti et al., and the curve of the vdW-dZK model for G-MoS2-4L is even similar to the ones obtained from the MBD (many-body dispersion) model [172].

The extended vdW-dZK method goes beyond the conventional LZK for treatments of many-body effects. It includes three features:

1. The conventional LZK uses a kind of local form, in which the substrate response is approximated in terms of an average dielectric function, computed at wave vector $q = 0$. In the vdW-dZK, non-local terms are included. Those are terms with $C^{nl}_{4,n}$ and $C^{nl}_{5,n}$, see Eqs. 2.13 and 2.14. The nonlocal terms C^{nl}_4 and C^{nl}_5 are derived from the hydrodynamic model at electrostatic limits, and are actually dipolar contributions contributing to the terms $\sim 1/D^4$ and $\sim 1/D^5$. In the derivations of the C^{nl}_4 and C^{nl}_5 terms, the wave vector $(q \neq 0)$ dependence was already considered, see details in [134, 135]. Those terms include effects of a nonlocal dielectric function.

2. The conventional LZK [130] was formulated for a particle interacting a semi-infinite surface. In our vdW-dZK, a finite-thickness surface (or a slab) is used and the image charge effects are accounted for explicitly. The image charge effects are also important for substrates of low dimensionality.

3. Our vdW-dZK includes the quadrupole term C^q_5. The quadrupole term contribution can be non-negligible. This was also previously found for graphene adsorbed on infinitely thick metal surfaces, see [136].

The synergistic effects of the above features in the vdW-dZK method not only produce reliable results for equilibrium configurations, but also give long-range power law variations of interlayer interactions.

CHAPTER 3

VAN DER WAALS CORRECTED DENSITY FUNCTIONALS FOR CYLINDRICAL SURFACES: AMMONIA AND NITROGEN DIOXIDE ADSORBED ON A SINGLE-WALLED CARBON NANOTUBE

3.1 Introduction

van der Waals (vdW) interactions are considered to be only a small contribution to the total energy [173, 174]. However, they are a key factor in describing the binding properties in a majority of molecular systems and materials. Although vdW interaction generally includes all intermolecular interactions, it more specifically (and everywhere here) refers to the London dispersion interaction [175, 176], a universal interaction between any pair of objects made of electrons. vdW interactions are strongly nonlocal, extending to distances exceeding $10nm$ (or $100\mathring{A}$) in nanoscale materials [177]. London-dispersion-based vdW interaction can be negligible in most solid systems; however, it has non-negligible effects in alkali metals [178], and plays a significant role in molecular complexes, molecular crystals, layered materials, surface adsorptions, and many biosystems. Its influence ranges from drug binding in proteins and double-helix stability in DNA [179] to pedal

adhesion in geckos [180, 181] and cohesion in asteroids [182, 183]. While the importance of understanding and modeling vdW interactions in realistic systems can hardly be overemphasized, our ability to accurately model vdW interactions from first principles is severely impeded by the high computational cost of the high-level methods. The adiabatic connection fluctuation-dissipation (ACFD) theorem [184, 185], or (after simplifying the interacting density response function) the random phase approximation (RPA) [186, 187], plays a central role in understanding the exact treatment of vdW interactions. However, the literature clearly lends support to the view that the ACFD or RPA based methods are not computationally efficient for large systems with over several hundreds of atoms per supercell. Therefore, many more efficient approaches (DFT-vdW) have been developed within the framework of density functional theory (DFT), as long-range additive corrections to semilocal approximations to the density functional for the exchange-correlation energy. These include the vdW-DF family [188, 189, 190, 191, 192, 193, 62], VV10 (Vydrov and Van Voorhis) [63] and rVV10 [194], DFT+D series [195, 196], and Tkatchenko-Scheffler (TS) methods [197, 198, 199]. These broadly support the view that DFT-vdW combined methods are perhaps one of the most efficient ways to include vdW interactions.

Carbon is a remarkable element showing a variety of stable forms [200] ranging from 3D semiconducting diamond to 2D semimetallic graphite to 1D conducting and semiconducting carbon nanotubes to 0D fullerenes, which show many interesting properties. Single-walled carbon nanotubes [201] (CNTs) are promising materials and building blocks for future nanoelectronics [202, 203, 204], due to their large surface area to volume ratios, unique electronic properties, as well as high chemical and mechanical stabilities.

Kong et al. [205] established a connection between the electrical conductivity of single-walled CNTs and their exposure to gaseous molecules such as O=N−O and NH_3. Inspired by their work, Chang et. al. [206] carried out a detailed study of the adsorption of NH_3 and NO_2 on CNTs. They emphasized designing the geometry and computing the binding energy, equilibrium distances, and charge transfers. However, they only took the local density functional (LDA) [35] into consideration and implemented double-numerical basis functions (DND) to compute the binding energies and equilibrium distances. Unfortunately, LDA does not correctly capture the medium and long-range interactions, and therefore the functional approximations needed to be modified. This realization motivated us to investigate the binding energies and equilibrium distances using GGA and meta-GGA functionals, i.e., Perdew-Burke-Ernzerhof (PBE) [40] and the strongly constrained and appropriately normed (SCAN) meta-generalized gradient approximation (meta-GGA) [207].

A natural question arises regarding the vdW interaction relating to nanotubes: to what extent does the vdW energy expression change from a flat to a cylindrical surface? Emig et al. [208], starting from the path integral representation, derived the exact Casimir

force between a plate and a cylinder. They found the force has an unexpectedly weak decay at large plate-cylinder separations, due to transverse magnetic modes. Rajter et al. [209] derived the formulas of vdW interactions for plate-cylinder, cylinder-cylinder systems, based on the Lifshitz theory. Starting from the vdW interaction between two separated semi-infinite half spaces, they then treated the half space as a composite of cylinder bundles and derived the vdW interactions for plate-cylinder and cylinder-cylinder systems. Their formulas can apply to systems made of metallic, semiconducting, and dielectrically isotropic or anisotropic materials, such as semiconducting or metallic carbon nanotubes. In general, vdW interactions are remarkably sensitive to the geometry and electronic structure of a given system. In this work, by using classical electrostatics starting from a dipole outside a cylindrical surface, we put forth a derivation of an expression for vdW energy between a particle and a cylinder. The particle-cylinder system is a prototype for molecules adsorbed to nanotubes. By using the developed formula, we compute the binding energies and equilibrium distances using our vdW model for NO_2 and NH_3 molecules adsorbed on CNTs. The present work is an extension of our developed damped Zaremba-Kohn model (vdW-dZK) model, which starts from a formula for the vdW interaction of a distant atom with a planar solid surface [210], both with known properties, damps this formula at short range, and then treats an adsorbed molecule or atomic layer as a collection of renormalized atoms. The vdW-dZK model has previously been successfully applied to study the physisorption of graphene and thiophene on metals [211, 212, 167] and graphene adsorbed on layered materials [213]. We also compare our results with GGA and meta-GGA combined with rVV10 [63, 194]. In the present work as in other DFT-vdW approaches, the semilocal density functional provides the short- to intermediate range interactions, while the fully nonlocal vdW term provides the intermediate- to long-range interactions, each term doing what it can do best.

3.2 Theoretical Methods

The adsorbed molecules we study here have permanent dipole moments that interact with the static charge densities they induce in the carbon nanotube substrate. That effect is already correctly included via the Hartree electrostatic energy in the approximate density functional calculations we present here. But time-dependent fluctuations of the electron densities around their static averages give rise to fluctuating dipole moments whose interaction creates the long-range dispersion attraction that is missing from semilocal approximations to the density functional for the exchange-correlation energy. Just as the Zaremba-Kohn treatment of the dispersion interaction starts from the interaction of a static dipole with a planar surface, we will start here from the interaction of a static dipole with

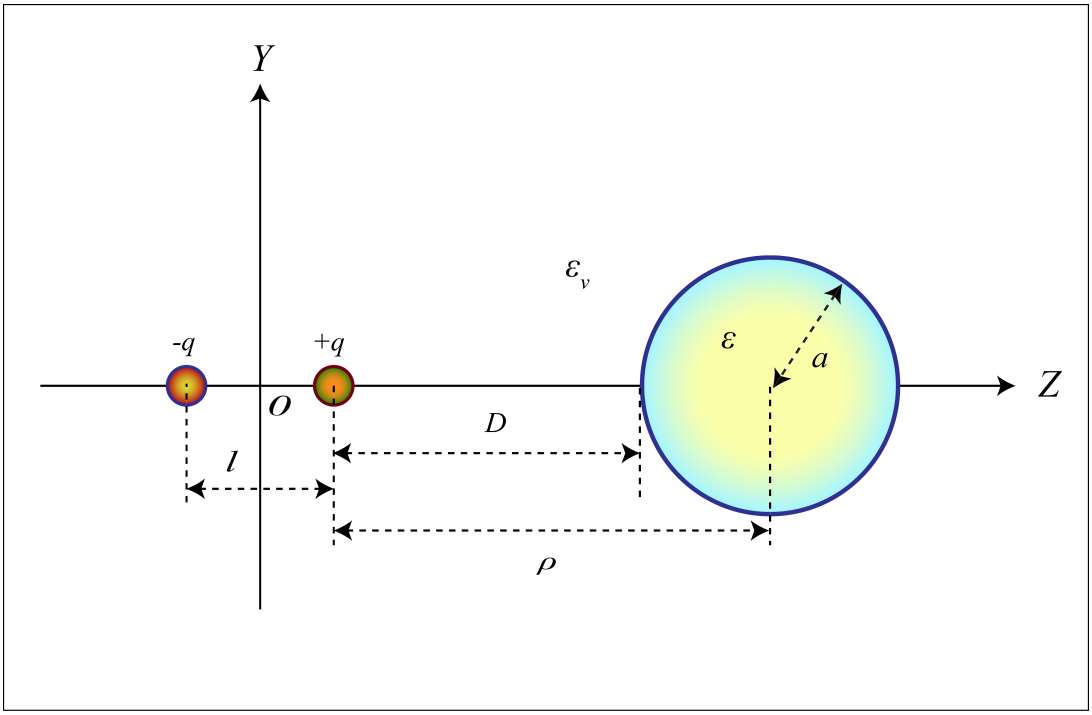

Figure 3.1: Schematic diagram of an interacting system consisting of a dipole and a dielectric cylinder. The origin of the coordinate system is at the center of the dipole. The cylinder has a relative dielectric constant ε and a radius a. The relative dielectric constant of the surrounding vacuum is $\varepsilon_v = 1$

a curved cylindrical surface. (The planar surface will emerge as the infinite-radius limit of the cylindrical surface, to make contact with Zaremba-Kohn theory.)

We start by calculating a dipole moment p, $p = ql$, in front of a dielectric cylinder with a radius a. The distance from the positive charge to the surface of the cylinder is denoted D. Consequently, ρ is defined as $D+a$. The direction of the dipole is perpendicular to the longitudinal central axis of the cylinder and pointed towards the cylinder. The origin of the coordinate system is at the center of the dipole, as shown in 3.1. Note that the dipole moment p here should be understood as an instantaneous dipole moment, not a permanent one. This means that p (hence q and l) is time dependent. The average of this dipole moment over a long period of time should be zero. However, the time average of p^2 (or q^2l^2) is not zero and will give the dispersion interaction. Since the vdW interaction is the interaction between fluctuating dipoles (or multipoles) at the electrostatic limit, we start from calculating the electrostatic interaction energy between a static dipole and the cylindrical surface.

The electrostatic interaction energy, E_{dipole}^{static}, between the dipole moment $p = ql$ and the charge density it induces in the dielectric cylinder can be calculated as a sum of

two parts E_1 and E_2, where the first part E_1 corresponds to the total electrostatic potential energy of the positive charge in the dipole moment p interacting with all the image charges residing inside the cylinder, and similarly, the second part E_2 corresponds to that of the negative charge in the dipole moment p interacting with all the image charges inside the cylinder.

The electric potential at a position $\vec{r}$, produced solely by a point charge q located at a position $\vec{r'}$ outside of a cylinder, is given by the following Green's function,

$$G_0(\vec{r},\vec{r'}) = \frac{q}{\varepsilon_v \left|\vec{r}-\vec{r'}\right|} \tag{3.1}$$

where ε_v is the relative dielectric constant of the vacuum surrounding the cylinder ($\varepsilon_v = 1$). All units are in atomic units unless otherwise specified. Eq. 3.1 can be expanded in cylindrical coordinates (ρ, z, ϕ), where the z axis is the central axis of the cylinder [214, 215], as

$$\begin{aligned} G_0\left(\vec{r},\vec{r'}\right) = \frac{4q}{\pi\varepsilon_v}\int_0^\infty dk\, cosk\left(z-z'\right)\Bigg[&\frac{1}{2}I_0\left(k\rho_<\right)K_0\left(k\rho_>\right) \\ &+\sum_{m=1}^{\infty} I_0\left(k\rho_<\right)K_0\left(k\rho_>\right) cosm\left(\phi-\phi'\right)\end{aligned} \tag{3.2}$$

where I_m and K_m are the modified Bessel functions of m-th order. These functions are defined, and their limits at small and large argument are presented, in Ref. [215]. $\rho_<$ or $\rho_>$ indicates the smaller or larger radial coordinate of the source and field points. k denotes a variable in the integration. ϕ and $\phi^{'}$ are the azimuth angles relating to the field and source points, respectively. z and $z^{'}$ are the z-coordinates of the field and source points, respectively. Using the above result and applying boundary conditions, the general form of the electric potential outside the cylinder, produced only by the image charges residing inside the cylindrical medium, is given by (see Eq. (30) of Ref. [215]),

$$G_0\left(\vec{r},\vec{r'}\right) = \frac{-2}{\pi\varepsilon_v}\int_0^\infty dk\, cosk\left(z-z'\right)\left[\sum_{m=-\infty}^{\infty} q_m(k)\frac{I_m\left(ka\right)}{K_m\left(ka\right)}I_m\left(k\rho_<\right)K_m\left(k\rho_>\right) cosm\left(\phi-\phi'\right)\right] \tag{3.3}$$

Here a charge q at position $r^{'}$ outside the cylinder produces image charges inside the cylinder which in turn create a potential at r outside the cylinder. In Eq. 3.3, $q_m(k)$ is a constant to satisfy the boundary conditions. It is expressed as,

$$q_m(k) = \frac{q(\varepsilon-\varepsilon_v)}{\varepsilon-\varepsilon_v h_m(k)} \quad \text{with} \quad m = 0, \pm1, \pm2, ..., \pm\infty \tag{3.4}$$

$$h_m(k) = \frac{I_m\left(ka\right)K_m^{'}\left(ka\right)}{K_m\left(ka\right)I_m^{'}\left(ka\right)} \tag{3.5}$$

where ε is the relative dielectric constant of the cylinder. From Eq. 3.2 to Eq. 3.3, the symmetry properties of the modified Bessel function (i.e., $I_{-m}(x) = I_m(x)$ and $K_{-m}(x) = K_m(x)$) are used. Please note that Eqs. 3.2 and 3.3 are expressed in a cylindrical coordinate system as in Ref. [215]. This is different from the coordinate system in 3.1. However, $\rho_<$ or $\rho_>$ in Eqs. 3.2 and 3.3 is related to ρ (ρ_+ or ρ_-) in 3.1. ρ_+ is the distance between the positive charge and the central axis of the cylinder and $\rho_+ = \rho = D + a$, and similarly, ρ_- is the distance between the negative charge and the central axis of the cylinder and $\rho_- = \rho + l = D + a + l$, as can be seen in 3.1. We will use Eq. 3.3 to calculate E^{static}_{dipole}, which includes E_1 and E_2. In 3.1, since the source or field point is either the positive or the negative charge, and the source and field points are on the same line perpendicular to the central axis of the cylinder, it follows that, in Eq. 3.3, we have $z = z'$ and $\phi = \phi^{'}$, and Eq. 3.3 is simplified as,

$$G_0\left(\vec{r},\vec{r'}\right) = \frac{-2}{\pi\varepsilon_v}\int_0^\infty dk\left[\sum_{m=-\infty}^{\infty} q_m(k)\frac{I_m(ka)}{K_m(ka)} I_m(k\rho_<)\,K_m(k\rho_>)\right] \tag{3.6}$$

In 3.1, the location of the positive charge in the dipole is $Z = l/2$, while that of the negative charge is $Z = -l/2$. With Eq. 3.6, the total electric potential $V_{Z=l/2}$ at the location of the positive charge of the dipole, produced by all image charges associated with both the positive and negative charges in the dipole, is expressed as,

$$\begin{aligned} V_{Z=l/2} = \frac{-2}{\pi}\int_0^\infty dk\,\Bigg\{ &\left[\sum_{m=-\infty}^{\infty} q_m(k)\frac{I_m(ka)}{K_m(ka)}\left[K_m(k\rho)\right]^2\right] \\ &+\left[\sum_{m=-\infty}^{\infty} q_m(k)\frac{I_m(ka)}{K_m(ka)}\left[K_m(k\rho)\,K_m(k\rho+kl)\right]\right] \end{aligned} \tag{3.7}$$

where the first square bracket term is the contribution from the image charges associated with the positive charge of the dipole and is obtained by letting both $\rho_<$ and $\rho_>$ equal ρ_+ (note $\rho_+ = \rho$) in Eq. 3.6. Note that for the first square bracket term, both the field point and the source point are at the location where the positive charge resides. The second bracket term in Eq. 3.7 is the contribution from the image charges associated with the negative charge and is obtained by letting $\rho_< = \rho_+ = \rho$ and $\rho_> = \rho_- = \rho + l$ in Eq. 3.6. Note that for the second square bracket term, the field point is at the location where the positive charge resides, and the source point is at the location where the negative charge resides. Because the second square bracket term in Eq. 3.7 is associated with the negative charge, there is a negative sign in front of $q_m(k)$. In the derivation of Eq. 3.7, $\varepsilon_v = 1$ is

used. So, we have,

$$E_1 = qV_{Z=l/2} = \frac{-2q}{\pi}\int_0^\infty dk \left\{ \left[\sum_{m=-\infty}^{\infty} q_m(k) \frac{I_m(ka)}{K_m(ka)} \left[K_m(k\rho)\right]^2 \right] + \left[\sum_{m=-\infty}^{\infty} q_m(k) \frac{I_m(ka)}{K_m(ka)} \left[K_m(k\rho)\, K_m(k\rho + kl)\right] \right] \right. \tag{3.8}$$

Similarly,

$$E_2 = -qV_{Z=-l/2} = \frac{-2q}{\pi}\int_0^\infty dk \left\{ \left[\sum_{m=-\infty}^{\infty} -q_m(k) \frac{I_m(ka)}{K_m(ka)} \left[K_m(k\rho + kl)\right]^2 \right] + \left[\sum_{m=-\infty}^{\infty} q_m(k) \frac{I_m(ka)}{K_m(ka)} \left[K_m(k\rho)\, K_m(k\rho + kl)\right] \right] \right. \tag{3.9}$$

The electrostatic interaction energy E_{dipole}^{static} is expressed as,

$$E_{dipole}^{static} = E_1 + E_2 = \frac{-2q}{\pi}\int_0^\infty dk \sum_{m=-\infty}^{\infty} q_m(k) \frac{I_m(ka)}{K_m(ka)} \left[K_m(k\rho)\, K_m(k\rho + kl)\right]^2 \tag{3.10}$$

which properly vanishes $\approx (ql)^2$ when $ql \to 0$. From this result, the total final vdW energy between the instantaneous dipole and the cylinder can be evaluated as (the full derivation can be found in the Appendix),

$$\boxed{E_{vdw} = -\frac{2}{\pi}\sum_{m=-\infty}^{\infty}\int_0^\infty dk \int_0^\infty du\; \alpha_1(iu)\xi_m(k,\rho)\Gamma_m(u,k) \left(\frac{D}{D+\bar{b}}\right)^2} \tag{3.11}$$

In Eq. (3.11), the damping function $\left(\frac{D}{D+\bar{b}}\right)^2$ is added to correct the spurious divergence as $D \to 0$. is a cutoff parameter and will be discussed in the following paragraphs. The function $\xi_m(k,\rho)$ and $\Gamma_m(u,k)$ are defined as,

$$\xi_m(k,\rho) = \frac{I_m(ka)}{K_m(ka)}\left[K_m'(k\rho)\, k\right]^2 \tag{3.12}$$

$$\Gamma_m(u,k) = \frac{\varepsilon(iu) - 1}{\varepsilon(iu) - h_m(k)} \tag{3.13}$$

and $\alpha_1(0), \bar{b}, \omega_1,$ and ω_p will be defined in the next two paragraphs. Eq. (3.11), along with Eqs. 3.5, 3.12, and 3.13, are the working equations for our calculations. As will be shown in the Results and Discussion section, Eq. (3.11) varies as $D^{-4.5}$ around $D = 50\mathring{A}$, and as D^0 when $D \to 0$.

A (n,m) CNT can be semiconducting or metallic, dependent on tube chirality. The armchair (n,n) CNTs are usually metallic and their longitudinal polarizabilities or dielectric constants are extremely large [216, 217] whereas the (n,0) zigzag CNTs can be

semiconducting or semimetallic. The semiconducting and metallic nanotubes have similar dielectric responses in the direction perpendicular to the tube axis, while their responses along the tube axis are very different [217]. In principle, our method can apply to any nanotube, whether metallic or semiconducting, as long as the dielectric function of the tube can be properly approximated. As a simple model, the dielectric response of a metallic nanotube can be modeled as that of a free-electron gas. Since the nanotube considered in this work is semiconducting, we use the modified Penn model to express the dynamic dielectric function, $\varepsilon(iu)$, of the semiconducting cylinder [142, 218]. It reads,

$$\begin{aligned}\varepsilon(iu) = 1 + \frac{\omega_p^2}{u^2}\left[\frac{\left(1-\Delta^2\right)y}{P} - \frac{\omega_g^2 - \left(\omega_g^2+u^2\right)\Delta^2}{2u\sqrt{\left(\omega_g^2+u^2\right)}} log\frac{I_+}{I_-}\right] \\ + \frac{2\omega_p^2\Delta}{u^2}\left\{\frac{\omega_g}{u}\left[tan^{-1}\left(\frac{\omega_g P}{u}\right) - tan^{-1}\left(\frac{\omega_g}{u}\right)\right] + \frac{1}{P} - 1\right\}\end{aligned} \tag{3.14}$$

where $I_\pm = \left[\left(1+y^2\right)\left(1+u^2/\omega_g^2\right)\right]^{1/2}, y = 1/\Delta$, and $P = \sqrt{1+y^2}$ ω_g is the effective energy gap, which is defined below, and $\Delta = \omega_g/(4\varepsilon_F)$ with $\varepsilon_F = (3\pi^2 n_{nanotube}^{2/3})/2$, with $n_{nanotube}$ defined in the second paragraph after Eq. 3.15. This model dielectric function has been used to study physical adsorption [142, 143, 144]. The effective energy gap ω_g is related to the static dielectric function $\varepsilon(0)$ via the Penn model [145],

$$\varepsilon(0) = 1 + \frac{2\omega_p^2}{3\omega_g^2}[(1+\Delta^2)^{1/2} - \Delta] \tag{3.15}$$

Since the accurate $\varepsilon(0)$ from a high-level *ab initio* calculation contains important hybridization, inhomogeneity, and exciton effects, which can be carried over to our model via the Penn model, the model dielectric function is expected to work well for semiconductors [144]. For a (10,0) semiconducting CNT, the average static dielectric function is evaluated as $\varepsilon(0) = 12.91$ from the data in ref. [217, 219]. ω_g is determined to be 0.289 atomic unit by Eq. (3.15). The carbon nanotube is treated here as a solid dielectric cylinder. We expect little error from this treatment, since (a) the interaction considered here is between the nanotube and an adsorbate outside the nanotube, and (b) the dielectric screening within the nanotube has been included to some extent in the effective dielectric constants. In the derivation of Eq. (3.11), the case of a cylinder with a large radius a is considered and an analogy to the flat-surface Zaremba-Kohn formula is also made (details see Appendix). This leads to Eq. (3.11) having features of the Zaremba-Kohn formula. However, Eq. (3.11) is a modification to a curved cylindrical surface. We call our model vdW-dZK (damped Zaremba-Kohn). Based on the estimated reference values (see Table 1), the cutoff parameter $\bar{b}$ is fitted by optimizing the mean errors (ME), root mean squared errors (RMSE), mean absolute errors (MAE), and mean absolute percentage errors (MAPE), and

consequently chosen by an "eyeball estimation" to be 1.7 Bohr (0.899 angstrom) for the PBE functional and 4.5 Bohr (2.38 angstrom) for the SCAN functional. A larger $\bar{b}$ results in a damping factor that turns on the long-range dispersion only at larger D. Note that the cutoff parameter $\bar{b}$ is relatively large for SCAN, due to the capture of intermediate correlations by SCAN. This is consistent with our previous work on layered materials, where $\bar{b}$ was relatively larger for SCAN-vdW-dZK than PBE-vdW-dZK.

The diameter of a (n,m) nanotube can be found as $d = \frac{\sqrt{3}a_{C-C}}{\pi}\sqrt{m^2+mn+n^2}$[201], where a_{C-C} is the nearest C-C distance $(1.42\mathring{A})$ [201]. Accordingly, the radius $a = d/2$ of the (10,0) nanotube is 7.398 Bohr $(3.915\mathring{A})$. We treat NH_3 and NO_2 molecules interacting with a nanotube as a collection of renormalized atoms (see Ref. [167] for details). Each atom has its renormalized static polarizability $\alpha_1(0)$ [167, 220]. For example, the renormalized static polarizability is 5.034 atomic units for N and 3.052 atomic units for H in the NH_3 molecule, while it is 7.988 atomic units for N and 5.824 atomic units for O in the NO_2 molecule. The dynamic polarizability of each renormalized atom is calculated as $\alpha_1(iu) = \alpha_1(0)\frac{\omega_1^2}{(\omega_1^2+u^2)}$ with $\omega_1 = \sqrt{\frac{4\pi n_{molecule}}{3}}$. where $n_{molecule}$ is the average valence electron density within the molecule [212, 167]. Furthermore, $\omega_p = \sqrt{4\pi n_{nanotube}}$. The carbon nanotube can be formed by rolling up a graphene sheet. The average electron density of the carbon nanotube can be approximated as that of the graphene sheet and is $n_{nanotube} = 8/[(2.46/0.529)^2 \times \sqrt{3}/2 \times 3.4] = 0.126$ atomic unit [211, 212].

Although individual factors in the integrand of Eq. (3.11) are highly divergent as $k \to 0$, the integrand itself has only a weak $1/log\ k$ divergence as$k \to 0$, and decays like $e^{-2k(\rho-a)}$ as $k \to \infty$. In the evaluation of Eq. (3.11), the integral over u is performed first; followed by the integral over k and then the summation over m. It is found that for distances D ranging from 2.1 to 10.1 Bohr $(1.11\mathring{A}$ to $5.34\mathring{A})$, the results obtained from the k grid of (0.001, 40, 0.001) and the m range of (0, 20, 1) deviate from those from the k grid of (0.001, 46, 0.001) and the m range of (0, 40, 1) by less than 0.03%. Note that we adopted the notation $(x_{min}, x_{max}, x_{increment}$ to describe the computation grids where x_{min} is the starting point, x_{max} is the ending point, and $x_{increments}$ is the increment. Within the same distance range, the results from the k grid of (0.001, 40, 0.001) and the m range of (0, 20, 1) are the same as those from the k grid of (0.001, 20, 0.001) and the m range of (0, 20, 1). Furthermore, within the same distance range, when using a denser k grid, the results from the k grid of (0.0001, 40, 0.0001) and the m range of (0, 20, 1) are different from those from the k grid of (0.001, 40, 0.001) and the m range of (0, 20, 1) by less than 0.2%. The above tests are done for the u grid of (0.01, 30, 0.01). At the k grid of (0.001, 40, 0.001) and the m range of (0, 20, 1), when using the u grid of (0.001, 50, 0.001), the deviation is less than 0.07%. For computational efficiency, we use the k grid of (0.001, 40, 0.001), the m range of (0, 20, 1) and the u grid of (0.01, 30, 0.01) for the rest of the calculations.

Since the equilibrium distances of the systems studied here are within this distance range, the results obtained here show acceptable convergence.

The form of the vdW-dZK Eq. (3.11) for a conducting cylindrical surface is very different from the expressions we obtained before for a flat metallic surface [211, 212] or a flat semiconducting layered material surface [213], although they all bear the same physics of the Zaremba-Kohn formula [210]. Only the dipolar effect is included in Eq. (3.11). High order terms, such as quadrupole terms, are very complicated and not included. Similar to the formula for a finite thickness semiconducting layered material surface slab, the multiple image effects are included in Eq. (3.11), however, through a complicated form with the modified Bessel functions. The parameter m can be regarded as an image index. However, the complicated form in Eq. (3.11) reflects the complexity of the shape and distribution of the images in the cylinder.

3.3 Results and Discussion

The (10,0) CNT is shown in Fig. 3.2. It is modelled by making a supercell (enclosed by faint lines in Fig. 3.2) of eight layers of carbon along the tube axis so that 80 atoms are included in the cell. The size of the supercell is about $28 \times 30 \times 8\,\mathring{A}$. The CNT is constructed with a nearest carbon-carbon distance of $1.42\,\mathring{A}$, then one NH_3 molecule is placed near the tube wall. Two rotational geometries are considered, as shown in Figs. 3.2(a) and 3.2(b) denoted type I and type II. The setup for the CNT + NO_2 system is similar. Two geometries of CNT + NO_2 are also considered, type I and type II, as shown in Figs. 3.2(c) and 3.2(d), respectively.

All DFT calculations are done in the Vienna Ab initio Software Package (VASP) [220] with projector augmented wave (PAW) pseudopotentials [221, 222]. The $1.42\,\mathring{A}$ carbon-carbon nearest-neighbor distance is a starting value and the nucleus of the N atom (or O atom) of the molecule is initially placed $3\,\mathring{A}$ straight above the nearest carbon atom of the tube wall, see Fig. 3.2. First, the molecule-tube structure is relaxed. Then, from the relaxed structure, the distance between molecule and tube is varied and the DFT binding energies at different distances are calculated so that the DFT curves of binding energy vs. distance are obtained. For the vdW energy calculations of the renormalized N atom (or O atom) in the molecule, the distance D is the distance between the nucleus of the N atom (or O atom) in the molecule and that of the nearest carbon atom. For the renormalized H atoms in the molecules, D is the distance between the nucleus of the H atom and the surface of a cylinder with the radius of the carbon nanotube. The energy cutoff is 580 eV. The k-point mesh is $1 \times 1 \times 3$. The structure optimization is conducted with all forces less than $0.01 eV/\mathring{A}$. The binding energy is defined as the total energy of the molecule-

nanotube structure after subtraction of the energy of the separated CNT and the energy of the separated molecule. Several Python scripts are written to compute the vdW energies using Eq. (3.11) for different distances, and the results are then combined with the DFT results to get the total binding energy curve, whose minimum is the equilibrium point, hence the equilibrium binding energy Eb and the equilibrium distance D_{eq}. The flow chart of the computational process is shown in Fig. 3.3. For comparison, the results from the SCAN+rVV10 and PBE+rVV10 calculations are also presented in Table 3.1.

Table 3.1: The binding energies $E_b$meV and equilibrium distances D_{eq} $\AA$ calculated from different methods for NH_3 and NO_2 molecules adsorbed on a 100 single-wall carbon nanotube

	Ref.	SCAN		SCAN+rVV10		SCAN+vdW-dZK		PBE		PBE+rVV10		PBE+vdW-dZK	
	E_b	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}	E_b	D_{eq}
Type I CNT+NH_3	-100	-43.3	3.38	-80.5	3.27	-103.4	3.22	-18.5	3.77	-70.8	3.44	-114.2	3.26
Type II CNT+NH_3	-100	-47.7	3.39	-84.2	3.31	-107.8	3.22	-18.4	3.77	-70.48	3.45	-113.8	3.26
Type I CNT+NO_2	-500	-354.3	2.70	-423.2	2.64	-486.0	2.58	-211.1	3.06	-309.3	2.80	-441.3	2.59
Type II CNT+NO_2	-500	-339.2	2.76	-402.7	2.69	-462.4	2.62	-207.7	3.16	-296.8	2.86	-422.3	2.62

As can be seen from Table 3.1, PBE captures no long-range and very little intermediate-range vdW interaction and yields an unphysically weak binding energy (-18.5 meV for CNT + NH_3 and -210 meV for CNT + NO_2). Furthermore, it results in a relatively large equilibrium distance of about 3.77 $\AA$ for CNT + NH_3 and about 3.06 $\AA$ for CNT + NO_2 . SCAN is designed to satisfy all 17 exact constraints that a meta-GGA functional can satisfy and can capture intermediate range (on the length scale of about 3 $\AA$) vdW interactions about right. Thus, SCAN gives a better description than does PBE. SCAN yields about half of the binding energies, compared with the reference, and shorter equilibrium distances than PBE does. Since it has been proven that SCAN is much more accurate than previous semilocal DFT methods, especially for geometrical and mechanical properties of materials, the equilibrium distances from SCAN should be closer to the real ones than those from PBE are. This is also consistent with the results we obtained in our layered material work [213].

As semilocal functionals, neither PBE nor SCAN include long-range vdW interactions. By adding the long-range vdW corrections (rVV10 or vdW-dZK) to PBE or SCAN, much better binding energies are achieved. This shows the significance of vdW energy in computing the total binding energy in the systems considered here. However, since the molecule-nanotube systems considered here are relatively large ($>$ 80 atoms), there is no high-level computational result in the literature, to the authors' best knowledge. The available binding energies E_b and equilibrium distance D_{eq} for NH_3 and NO_2 molecules adsorbed on carbon nanotubes from the literature are calculated with LDA, PBE+D, or

vdW-DF functionals. For CNT+NH_3 , Ref. [61] gives E_b = -0.15eV and D_{eq} = 2.99 $\AA$ with LDA. Ref. [223] gives E_b = -0.14eV and D_{eq} = 3.08 with LDA. Ref. [206] gives E_b = -0.18eV and D_{eq} = 2.90 $\AA$ with LDA. However, Ref. [224] gives E_b = -0.043eV and D_{eq} = 3.50 $\AA$ with vdW-DF. For CNT+NO_2 , Refs. [206, 225, 226] show that the binding energy ranges from -0.34 to -0.79 eV, and the equilibrium distance ranges from 1.93 to 2.61 $\AA$. However, Ref. [224] gives E_b = -0.26eV and D_{eq} = 3.20 $\AA$ with PBE+D. Based on these data, we estimate average values of binding energies for the two systems. The estimated values of binding energies are also listed in Table 3.1. The estimated references listed here can only serve as a rough guideline. Nevertheless, the results shown in Table 3.1 give a clear trend, from which we can find some physically relevant clues.

For NH_3 adsorbed on the carbon nanotube, the two adsorption configurations (Types I and II) have very similar binding energies. The predicted values of the binding energies using SCAN, SCAN+rVV10, and SCAN-vdW-dZK are slightly smaller for Type I than for Type II, while the predicted values of the binding energies using PBE, PBE+rVV10 and PBE-vdW-dZK are slightly larger for Type I than for Type II. For NO_2 adsorbed on the nanotube, the predicted binding energies (about 200-500meV) are higher than the case of CNT+NH_3 . It was reported [206] that CNTs respond more quickly and sensitively to NO_2 gas than to NH_3 gas. This is clearly illustrated by stronger binding energy from NO_2 than from NH_3 . All methods in Table 3.1 predict that the binding energies from type I are slightly larger than those from type II for CNT+NO_2 . However, the differences in binding energies between the two configurations for both CNT+NH_3 and CNT+NO_2 are very small, indicating similar binding stability between the two configurations [206]. Moreover, the values of the binding energies are relatively small ($\leq$ 500 meV) for both CNT+NH_3 and CNT+NO_2 , indicating that the adsorption of the two molecules on nanotubes is physisorption.

With the combinations of the long-range vdW functional rVV10, both PBE and SCAN significantly improve in their descriptions for binding energies. However, PBE+rVV10 produces about 15% and 25% lower binding energies than SCAN+rVV10 does for CNT+NH_3 and CNT+NO_2 , respectively. SCAN+rVV10 delivers a generally improved description for many systems. However, SCAN+rVV10 has some serious issues, including the overestimated equilibrium mass density in liquid water [227], inaccurate structural and mechanical properties in PPTA [228], overestimated binding energies of thiophene molecule adsorbed on metal surfaces [167] and graphene adsorbed on transition metal dichalcogenide multilayer materials [213], and inconsistent errors of overbinding and underbinding in different 2D materials [229]. The drawback of SCAN+rVV10 may be due to its imperfect balance of interactions at the intermediate range. At present, although the results of SCAN+rVV10 seem very reasonable, it is hard to determine its accuracy, due to the shortage of accurate

reference data. The same imperfect balance of interactions at the intermediate range also exists in PBE+rVV10. This leads to the limited accuracy of PBE+rVV10.

When combined with the developed vdW-dZK method, both PBE and SCAN improve in their descriptions of adsorption for CNT+NH_3 and CNT+NO_2 significantly. As can be seen in Table 3.1, SCAN-vdW-dZK and PBE-vdW-dZK give approximately the same results for CNT+NH_3 . For CNT+NO_2 , although PBE-vdW-dZK gives about 10% lower binding energies than does SCAN-vdW-dZK, the results of binding energies from SCAN-vdW-dZK and PBE-vdW-dZK are the closest to each other. This also shows the consistency of our developed SCAN-vdW-dZK and PBE-vdW-dZK methods. Based on this consistency, it seems that the results from SCAN-vdW-dZK and PBE-vdW-dZK are closer to the real ones. There are still gaps between the results of SCAN-vdW-dZK and PBE-vdW-dZK and the estimated reference in Table 3.1. However, since the estimated references are averages of limited data from the literature, the references present just a rough guideline, not an accurate one.

Ambrosetti et al. [229] have shown nontrivial variations of power-laws of the vdW interactions in systems of atoms or small molecules and 2D thin metallic or finite-gap substrates with RPA-like and CDH (coupled dipolar Hamiltonian) treatments for substrate responses. They provided evidence to support the claim that the power laws of vdW interactions substantially deviate from the standard pairwise predictions. The power-law exponent $P(D)$ is defined as $P(D) = d(log|E_{vdw}|)/dlogD$, and is a decaying function of separation D. We also calculate the vdW interaction energy power laws from our vdW-dZK model for the CNT+NH_3 and CNT+NO_2 systems. Fig. 3.4 shows the vdW power-law exponent from our vdW-dZK model as a function of D. It shows that the vdW-dZK exponent is zero at small D. The behavior of the power law exponent at small distance ($< 10\,\AA$) is more influenced by the damping function. At distances D about $20\,\AA$, the exponent is about -4.0 and at distances D about $50\ \AA$, the exponent is about -4.5, clearly keeping away from the pairwise limit -5. It is apparent that, at even larger distances, the vdW power-law exponent will go to -5. This is the expected value for the non-retarded vdW interaction of a molecule distant from a 1D insulator. This feature is very similar to the one calculated by Ambrosetti et al., see the second curve in the upper figure of Fig. 3.2 in Ref. [229], where the lattice constant of the carbon atom chain is $1.4\,\AA$, very close to that of the carbon nanotube. In Ref. [229], the results are obtained from RPA and RG (renormalization group) approaches. The similarity between our results and that of Ref. [229] provides support to our methods. The finite ($3.9\,\AA$) radius of the carbon nanotube and the damping factor are also at play in our results.

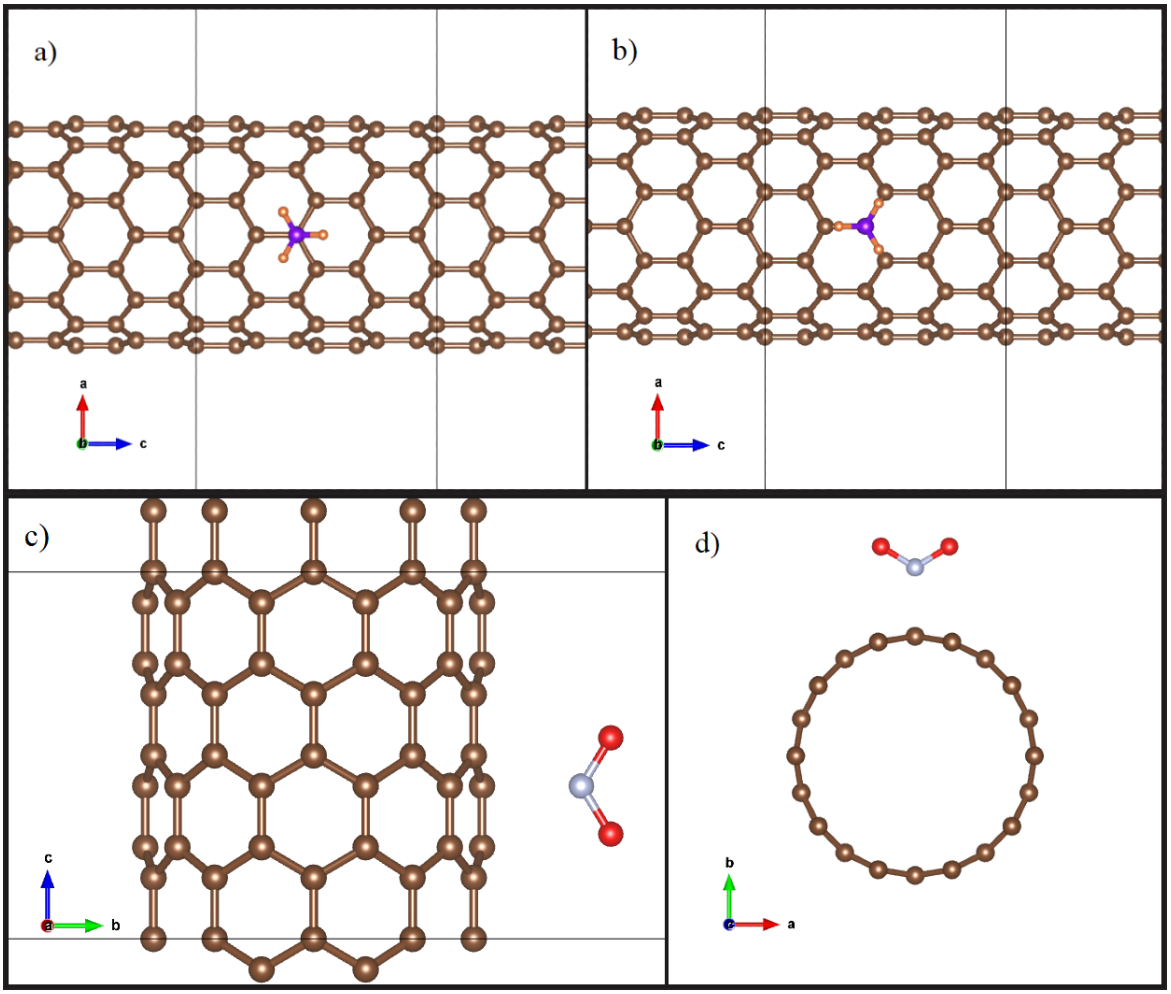

Figure 3.2: (10,0) CNTs with adsorbed NH_3 or NO_2 molecules. a) Top view of the NH_3 molecule attached to the CNT in Type I geometry. b) In Type II geometry. c) Side view of the NO_2 molecule attached to the CNT when the plane of the NO_2 is parallel to the axis of the CNT (Type I). d) Side view of the NO_2 molecule attached to the CNT when the plane of the NO_2 is perpendicular to the axis of the CNT (Type II).

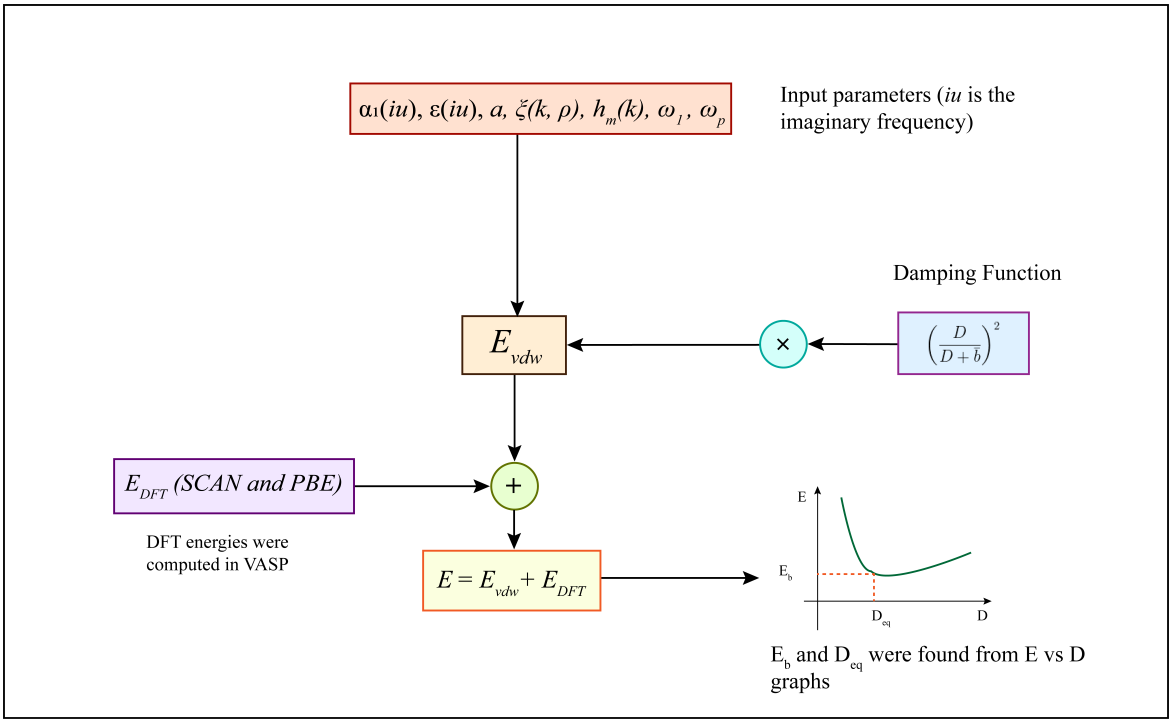

Figure 3.3: Flow chart of the computational process for the vdW interaction energies, DFT energies, and total energies for the developed PBE-vdW-dZK and SCAN-vdW-dZK methods.

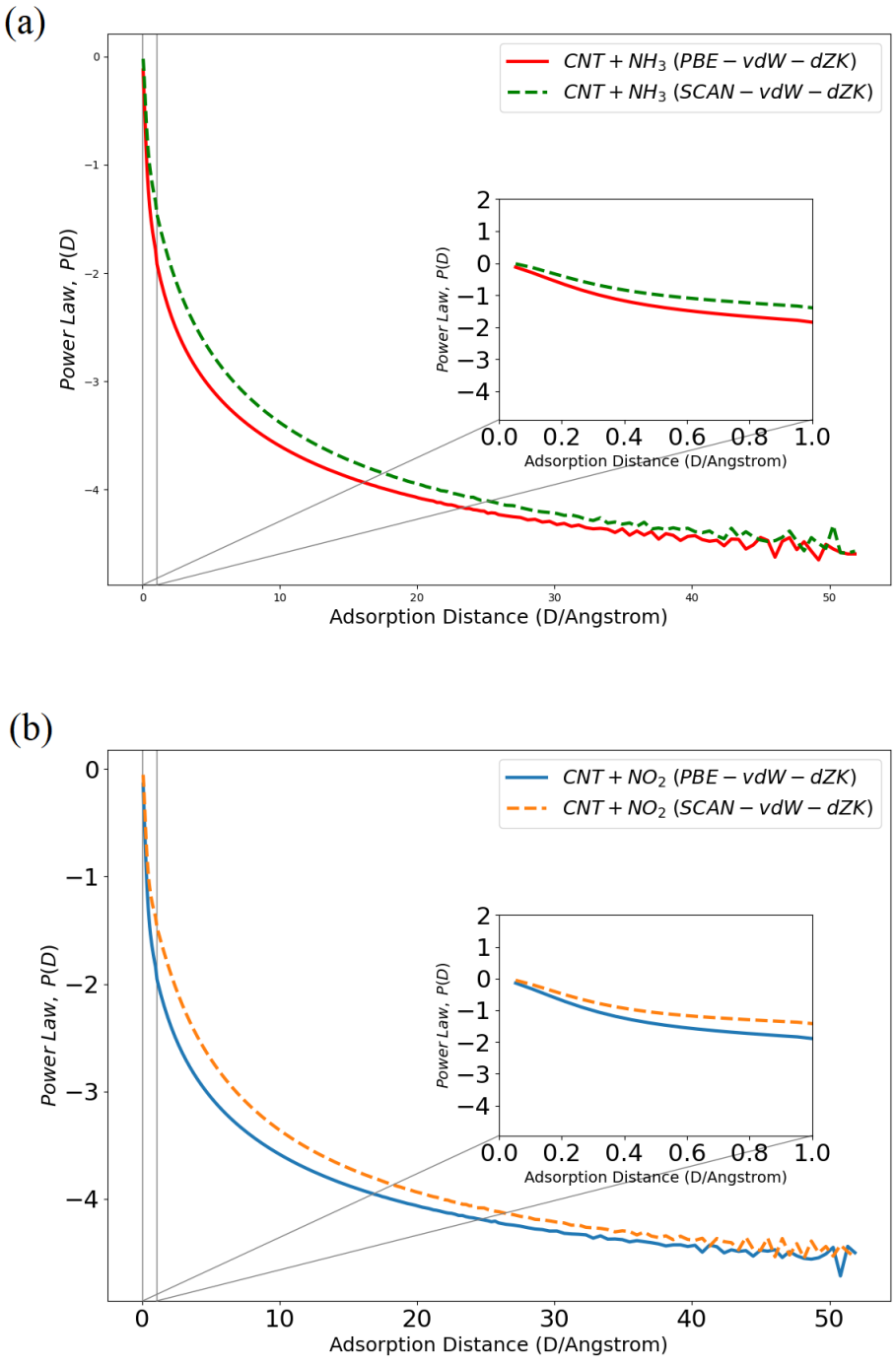

Figure 3.4: vdW interaction energy power law exponents from the vdW-dZK model for (a) NH3 and (b) NO_2 molecules adsorbed on CNT as a function of adsorption distance D.

CHAPTER 4

SPHERICAL VS. NON-SPHERICAL AND SYMMETRY-PRESERVING VS. SYMMETRY-BREAKING DENSITIES OF OPEN-SHELL ATOMS IN DENSITY FUNCTIONAL THEORY

4.1 Introduction

Density-functional theory (DFT) is one of the most popular and successful quantum mechanical approaches to ground-state matter. It is nowadays routinely applied to calculate, e.g., the structures and binding energies of molecules in chemistry and of solids in physics. The atomization energies of molecules, or the energies needed to break all the bonds between the atoms, have long been important tests of approximate density functionals. These tests are straightforward when all the atoms are closed-shell like He or closed-subshell like H, and when all the molecules are similarly closed-shell or closed sub-shell. When this is not the case, the standard self-consistent Kohn-Sham calculations require further justification (some of it to be provided here) beyond the level of their numerical success. In an open-subshell atom, the approximated Kohn-Sham one-electron potential can be non-spherical, unlike the spherically-symmetric external potential, leading to a self-consistent electron density with spatial symmetry lower than that of the exact (non-spherical) density. That is the only symmetry breaking discussed in most of this article. While symmetries can break over long time intervals in reality, making symmetry breaking real or at least revealing, a single main-group atom is expected to be too small and too normally correlated

to exhibit real symmetry breaking [230, 231]. By imposing a partial (first-step) space- or space-spin symmetry preservation, and thus showing that the space- and space-spin symmetry-breaking by approximate density functionals is energetically small for open-shell atoms, we provide more justification for the use of atomization energies of molecules (and solids), as standardly calculated, to test the approximate functionals. Here "small" is in comparison to the energy change from a spherical (ensemble) density to a non-spherical one, and also in comparison to the error reductions in the sequence Hartree-Fock, LSDA, PBE, and SCAN.

In 1964 Hohenberg and Kohn [14] showed that there exists a universal non-relativistic density functional $F[n]$, independent of the external potential $v_{ext}(\mathbf{r})$ (e.g., the attraction of the electrons to the nuclei), such that minimization of the sum

$$F[n] + \int n(\mathbf{r})\, v_{ext}(\mathbf{r})\, d^3\mathbf{r} \tag{4.1}$$

subject to the constraint

$$\int n(\mathbf{r})\, d^3\mathbf{r} = N \tag{4.2}$$

yields the ground-state energy and electron density of a quantum-mechanical N-electron system moving in this external potential. The Hohenberg-Kohn existence theorem has motivated the search for practical approximations to $F[n]$. Kohn and Sham [19] showed that a large part of F[n] could be constructed from self-consistent one-electron wavefunctions or orbitals that are eigenstates of a self-consistent one-electron effective Hamiltonian, leaving only the density functional for the exchange-correlation energy to be approximated. The computational cost of a Kohn-Sham calculation is far less than that of a correlated-wavefunction calculation, especially for large N. Kohn-Sham spin-density functional theory [19] proved to be more accurate than Kohn-Sham total-density functional theory, because of the extra information that it provides to the approximate functionals. Here we will work within Kohn-Sham spin-density functional theory, with the further common assumption that there is no spin-dependence in the external potential or in the electron-electron interaction (which we take to be Coulombic).

The original proof of the Hohenberg-Kohn theorem was restricted to non-degenerate ground states [14], and the set of densities over which to minimize was restricted to the ground-state densities for the class of scalar external potentials $v_{ext}(\mathbf{r})$. These restrictions were removed in the Levy proof [15], which starts from the variational principle for the many-electron wavefunction, then performs the search over wavefunctions in two steps: First over all wavefunctions that yield a given density, and then over all densities that come from any wavefunction (i.e., over all reasonable densities). The Levy proof can be extended [15] from a constrained search over wavefunctions to a constrained search over ensembles, which yields the same ground-state energy but different density functionals and

an electron density with the full symmetry of the external potential. A remaining question for the exact Kohn-Sham theory (but not one we will much consider here) is whether the ground-state density of the real system can be replicated by the ground-state density of a system of non-interacting electrons in an effective scalar external potential $v_{ext}(\mathbf{r})$. The answer to this question of non-interacting v-representability is yes for the ensemble search, but not necessarily always yes for the wavefunction search [15]. It is possible that the Kohn-Sham wavefunction [19] of lowest interacting energy is a low-lying excited state of a non-interacting system. The wavefunction of the Kohn-Sham auxiliary system of non-interacting electrons is a single Slater determinant or a linear combination of a few such determinants that are degenerate at the non-interacting level. Here we will consider only atoms and molecules in equilibrium, but there are also interesting effects in the binding energy curves of molecules.

Janak et. al. [232] provided evidence that, within the local spin density approximation (LSDA) [19] for the exchange-correlation energy, non-spherical corrections to the energy are quite small in spin-polarized calculations for first-row atoms and can be accurately calculated by first-order perturbation theory for cases where the corrections are significant (transition-metal atoms and non-spin polarized treatments). Variational considerations lead one to expect that removing the constraint of spherical symmetry would lower the atomic energy. Fractional occupation numbers arise naturally in an ensemble picture. Kutzler and Painter [233] found much larger non-spherical corrections to the energy and strongly improved atomization energies for molecules from an early constraint-based generalized gradient approximation (GGA).

A local exchange-correlation energy density and potential at a point in space depends only on the electron spin densities at that point [14, 19]. In the past few years, advances have been made in the development of the computationally-semilocal GGA [13-17] and meta-GGA [40]. A semi-local exchange-correlation energy density depends not only on the density at the point of interest but also on the gradient of the density at that point, and possibly on further information such as the non-interacting positive kinetic energy density there. This development of density-functional theory improves the predicted binding energies of sp-bonded molecules [19]. The beyond-LSDA functionals considered here are the Perdew-Burke-Ernzerhof (PBE) [35] GGA and the strongly constrained and appropriately normed (SCAN) meta-GGA) [40]. Like LSDA, those functionals are constructed by the satisfaction of exact constraints and are not fitted to the properties of any bonded systems.

The paper seeks to address the following question: Should any symmetry be imposed on the electron densities of open-shell atoms in DFT? In the next section, we report results from several calculations with approximate density functionals which show that

the total energies of non-spherical atoms are systematically lower than those for spherical atoms, a result which leads to appreciably improved molecular binding energies. Next, we touch on the issue of the symmetry of the Hamiltonian and the ground-state density. Ref. [234] suggests that the Kohn-Sham noninteracting wave function need not display the symmetries of the interacting wave function. However, it must produce the correct spin densities, which are influenced by symmetry. While the ground state density has the full symmetry of the Hamiltonian in thermal-equilibrium ground ensembles and non-degenerate pure ground states, our work leads us to conclude that when there are degenerate pure ground states the best approximate functionals imitate the constrained search over pure states and not that over ensembles. The symmetry of the interacting ground-state wavefunction can be broken by the approximate Kohn-Sham non-interacting wavefunction, but the energetic consequences of that symmetry breaking in atoms as found here are too small to be important.

Our work extends the 1987 work of Kutzler and Painter [233] to the more modern PBE GGA [35] and to the SCAN meta-GGA [40]. Like Ref. [233], we focus on the atoms and homonuclear diatomic molecules of the atoms B, C, O, and F. Those atoms have open p subshells and non-spherical ground-state densities. The work of Ref. [233] was done at a time when fully self-consistent calculations (now the norm) were uncommon. Instead of performing fully self-consistent calculations, Ref. [233] used an equilibrium ensemble to make a Kohn-Sham effective potential with the same spatial symmetry as the external potential. For an atom, this approach yields a spherical density and a spherical Kohn-Sham effective potential. Ref. [233] then constructed a possibly less symmetrical density (e.g., a non-spherical density for an atom) by making integer occupations of the Kohn-Sham orbitals from that ensemble Kohn-Sham potential, with that potential having spherical symmetry for atoms and cylindrical symmetry for diatomic molecules. For many open-shell cases, this is not the self-consistent broken-symmetry solution that yields the lowest energy. We will start with their approach, calling it non-spherical Sym-P1 in section 2. Later we will recognize in it the first steps of a general approach to symmetry preservation in density functional theory, and compare it to the now-standard self-consistent approach that can break symmetry. We will first discuss the energetic effects of spatial symmetry breaking in the atoms, utilizing the computational approach of Kutzler and Painter [233], and later we will extend their approach to the energetic effects of spin symmetry breaking in the atoms.

4.2 Calculations using Spherical and Non-Spherical Atomic Densities

In this section, we assess the impact of spherical and non-spherical atomic densities on calculated energies, following the Kutzler-Painter [233] approach described at the end of the preceding section. Anticipating a later discussion, we will refer to their approach as first-step non-self-consistent space symmetry-preserving (Sym-P1). As in their work, real Cartesian orbitals are chosen (and spin symmetry breaking in closed sub-shell molecules is not allowed.) Table 4.1 presents the spherical and non-spherical energies for several functionals. In all cases, we construct the state of maximum possible z-component of total spin, which for non-interacting pure states we take to be a single Slater determinant [235]. The electron configurations in the valence-shell integer-occupation or pure-state scheme are p_z^1 for boron (B), $p_x^1\ p_y^1$ for carbon (C), $p_x^1\ p_y^1\ p_z^2$ for oxygen (O), and $p_x^2\ p_y^2\ p_z^1$ for fluorine (F). The corresponding atomic densities have cylindrical symmetry about the z-axis. Likewise, the electron configurations in the valence-shell fractional-occupation or equilibrium-ensemble scheme are $p_x^{1/3}\ p_y^{1/3}\ p_z^{1/3}$ for B, and $p_x^{2/3}\ p_y^{2/3}\ p_z^{2/3}$ for C, all with z-component of spin S_α . For O and F, perhaps it is more revealing to present the fractional occupations divided into α and β spin channels (S_α and S_β). For the oxygen atom,

$$S_\alpha\colon\ p_x^1\ p_y^1\ p_z^1$$
$$S_\beta\colon\ p_x^{1/3}\ p_y^{1/3}\ p_z^{1/3}$$

Likewise, for the fluorine atom,

$$S_\alpha\colon\ p_x^1\ p_y^1\ p_z^1$$
$$S_\beta\colon\ p_x^{2/3}\ p_y^{2/3}\ p_z^{2/3}$$

The corresponding atomic densities have spherical symmetry.

All DFT calculations for atoms and molecules were carried out in all-electron NWChem [236] using the unrestricted Kohn-Sham approach, allowing for a spin-dependent exchange-correlation potential. NWChem permits the use of fractional occupation numbers without additional coding. For a given spin multiplicity $2S+1$, the z-component of total spin was set to S. The basis set was 6-311++G (3df,3pd), which converges valence-electron energy differences in Kohn-Sham DFT. For atoms, the spherical potential from the fractional-occupation configurations was used to generate the integer-occupied p orbitals. In other words, the same p orbitals are used in the spherical (equilibrium-ensemble) and non-spherical (symmetry-preserving, as discussed later) calculations, and only the occupations are changed. The numerical integration necessary for the evaluation of the exchange-correlation energy implemented in NWChem uses an Euler-MacLaurin scheme for the radial components (with a modified Mura-Knowles transformation) and a Lebedev scheme

for the angular components. We use two levels of accuracy (the `xfine` and `huge` grids) for the numerical integration to get the total energy target accuracy of 1×10^{-8} and 1×10^{-10} Hartree. The biggest relative difference in atomization energies between these two target accuracies is only a quarter of a percent even for SCAN (see Appendix C). The maximum number of iterations is set to 100. It should be noted that SCAN, PBE and our LSDA agree exactly for all uniform spin densities. Our LSDA uses the exact exchange energy per electron of a uniform electron gas and the PW92 [237] parametrization of the correlation energy per electron for uniform spin densities.

Table 4.1: Effect of the removal of the spherical approximation on the atomic energies of B, C, O, and F with three nonempirical density functionals (SCAN, PBE, and LSDA). Sym-P1 stands for non-spherical (first-step non-self-consistent symmetry-preserving) and Spherical for spherical densities. Energies are in Hartree, unless otherwise specified. 1 Hartree = 27.21 eV.

Atom	**SCAN**	**PBE**	**PW92**
B (Spherical)	-24.6216	-24.6032	-24.3504
B (Sym-P1)	-24.6364	-24.6085	-24.3520
Diff (eV)	0.40	0.14	0.04
C (Spherical)	-37.8181	-37.7903	-37.4650
C (Sym-P1)	-37.8343	-37.7939	-37.4644
Diff (eV)	0.44	0.10	-0.02
O (Spherical)	-75.0355	-74.9933	-74.5173
O (Sym-P1)	-75.0620	-75.0041	-74.5188
Diff (eV)	0.72	0.29	0.04
F (Spherical)	-99.7047	-99.6542	-99.0998
F (Sym-P1)	-99.7328	-99.6613	-99.0979
Diff (eV)	0.76	0.19	-0.05

From Table 4.1, it is apparent that PBE and especially SCAN energies are lowered significantly when evaluated with non-spherical densities. Density sensitivity increases from LSDA to PBE to SCAN, as might have been expected from the fact that the LSDA exchange-correlation energy density at a position in space depends only on the local spin densities, while in PBE it depends also on the local density gradients and in SCAN it depends further on the local non-interacting positive kinetic energy density. As can be seen in Table I, within the PBE approximation oxygen shows the largest non-spherical effect, with the total energy in the non-spherical treatment lying 0.29 eV lower than the result in the spherical approximation. For the advanced semi-local functional SCAN, the

atomic energies of all four atoms are significantly lowered by including the non-spherical corrections (relaxing the spherical constraint). The largest difference is observed in the fluorine atom, where the SCAN functional gives an energy lowering of about 0.76 eV for the non-spherical atom compared with that in the spherical approximation, while the smallest effect occurs in the boron atom; 0.40 eV. The results for LSDA and GGA are in good agreement with those of Ref. [233]. Furthermore, using the same functionals as in Ref. [233], we reproduced similar energies (Appendix C). The similarity between our results and those of Ref. [233] provides support for the correctness of our computations.

Table 4.2: Binding energies of first-row dimers using spherical (equilibrium-ensemble) and non-spherical (first-step non-self-consistent space symmetry-preserving) atomic densities. These densities of the atoms are paired with the naturally cylindrical densities of the dimers. Energies in electron volts (eV). The reference atomization energies are those experimentally observed for the ground state [12,24]. (See Table III for the electronic configurations of the dimers.) MAE is mean absolute error, widely used in density functional theory and permitting comparison with Table IV. (To find the atomization energies from self-consistent calculations with SCAN, PBE, and LSDA, subtract twice the small "Difference" in Table V from the entry here for Sym-P1. The errors of the self-consistent atomization energies are displayed in Fig. 1.)

		Non-spherical			**Spherical**		
Dimer	**Reference**	**SCAN**	**PBE**	**LSDA**	**SCAN**	**PBE**	**LSDA**
B_2	3.01	3.18	3.40	3.88	3.98	3.69	3.97
C_2	6.22	5.37	6.31	7.29	6.25	6.51	7.26
O_2	5.12	5.76	6.39	7.69	7.20	6.98	7.78
F_2	1.60	1.81	2.42	3.46	3.34	2.81	3.36
MAE		0.47	0.64	1.59	1.21	1.01	1.61

Passing to the atomization energies of molecules, we see from Table 4.2 that inclusion of both nonlocal and non-spherical corrections (or relaxing the spherical constraint on the atomic densities) gives closer agreement between theoretical and experimental binding energies of the first-row dimers. The binding energies are calculated from the experimental ground-state configurations of the molecules (${}_3\Sigma_g^-, {}_1\Sigma_g^+, {}_3\Sigma_g^-, {}_1\Sigma_g^+$ for B_2, C_2, O_2 and F_2 respectively). Furthermore, the employed bond lengths of B_2, C_2, O_2 and F_2 are $1.59\mathring{A}, 1.243\mathring{A}, 1.208\mathring{A}$, and $1.412\mathring{A}$ respectively. Table 4.2 clearly illustrates that, while non-sphericity alone brings some improvement in calculated molecular binding energies, the use of nonlocal functionals in the atom calculations leads to significant further reduc-

tions in the errors. In fact, SCAN, when combined with a non-spherical density, produces the lowest mean absolute error or MAE (0.47 eV). In contrast, for a non-spherical density, the PBE functional yields an MAE of 0.64 eV and LSDA yields a large MAE of 1.59 eV. Furthermore, comparison between the non-spherical and spherical results for the same functional shows that non-spherical densities almost always result in an atomization energy closer to the experimental value. A possible inference from the results is that the sequence of approximate functionals LSDA, PBE, and SCAN is converging toward the exact density functional defined by a constrained search over wavefunctions, and not to the one defined by a constrained search over ensembles. One reason might be that the lower-symmetry densities of wavefunctions provide more information to the functional than do the higher-symmetry densities of some ensembles. A similar argument explains why approximate spin-density functionals are more accurate than approximate total-density functionals, even in systems where the external potential is spin-independent.

Fig. 4.1 shows the errors of self-consistently calculated atomization energies, as given by the Sym-P1 atomization energies of Table 4.2 minus twice the small "Difference" from Table 4.4. Note the improvement from LSDA to PBE to SCAN for the atomization energies in Fig. 4.1: While the change from LSDA to PBE is roughly a constant shift toward weaker bonding, the change from PBE to SCAN is more properly system-specific, except for C2 whose abnormality is discussed at the end of this section.

What is the effect of breaking the symmetry of the atoms and molecules considered here? The ground state electronic configurations of our molecules are illustrated in Table III. All the molecules studied here have $D_{\infty h}$ point symmetry group [1] and have cylindrically symmetric ground-state densities. Thus, they have the symmetry of the external potential. For them, there is no difference in our calculations among the ensemble cylindrically symmetric, the space symmetry-preserving, and the self-consistent densities.

The case of C_2 is, however, intriguing. The straightforward application of the *aufbau* or building-up principle suggests that the ground-state configuration of C_2 at the equilibrium geometry is a singlet configuration $1\sigma_g^2 1\sigma_u^2 1\pi_{ux}^2 1\pi_{uy}^2$ [238], which is also cylindrically symmetric. However, we are dealing with a many-electron molecule, and the occupation of the lowest energy orbitals does not necessarily lead to the lowest energy. There is a possibility that excitation of an electron to a nearby orbital might lower the electron–electron repulsion and result in a lower overall energy despite the occupation of a higher energy orbital. The resulting configuration is a triplet with configuration $1\sigma_g^2 1\sigma_u^2 1\pi_{ux}^2$ $1\pi_{uy}^1 2\sigma_g^1$ [239, 238], which is not cylindrically symmetric (for real orbitals). Therefore,

[1] In the Schoenflies system, the D_{nh} group consists of the operations present in D_n together with a horizontal reflection, in addition to whatever operations the presence of these operations implies. An important example is the $D_{\infty h}$ group, to which uniform cylinders and homonuclear diatomic molecules belong.

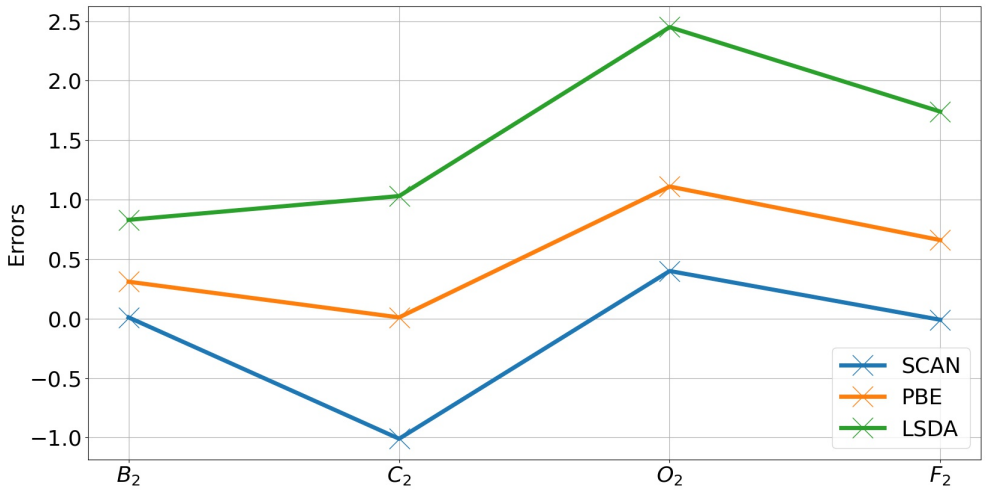

Figure 4.1: Errors (in electron volts) of the self-consistently calculated atomization energies of four diatomic molecules from the LSDA, PBE, and SCAN density functionals, in electron volts. The mean absolute errors (MAE) are 0.36 eV (SCAN), 0.52 eV (PBE), and 1.51 eV (LSDA). Spin symmetry breaking in strongly correlated C2 has been suppressed, but if included it would reduce the errors of PBE and SCAN.

Ref. [239] suggests that singlet and triplet states compete in energy for C_2. This competition seems to be confirmed in the CCSD(T) total energies with the cc-pCVTZ basis set reported in Ref. [24]. But the story of singlet C_2 is complicated even further by its multi-reference strong correlation attributed [240] to an avoided crossing between two states of the same symmetry near the equilibrium bond length. Only some of this strong correlation can be captured by SCAN, leading to an unusually strong underbinding of singlet C_2 when the spin symmetry of the molecule is not allowed to break. In the full configuration interaction quantum Monte Carlo calculation of Ref. [240], the more normally correlated triplet state is a very low-lying excitation, lying vertically above the singlet by only about 0.2 eV.

Spatial symmetry does not break in our approximate Kohn-Sham descriptions of our cylindrical molecules. Spatial symmetry can break in the non-spherical open-subshell atoms B, C, O, and F, but as we will see the energy consequences of that are small.

4.3 What Does the Symmetry of the Hamiltonian Say About The Symmetry of the Ground-State Density?

In this section, we summarize the known relationships between symmetry and degeneracy which are applied in this work. A symmetry operator $\hat{U}$ of a Hamiltonian operator $\hat{H}$ is a unitary operator that leaves the Hamiltonian invariant:

$$\hat{U}\hat{H}\hat{U}^{\dagger} = \hat{H} \tag{4.3}$$

Since $\hat{U}$ is unitary, $\hat{U}^{\dagger} = \hat{U}^{-1}$ and $\hat{U} = e^{i\hat{A}}$ where the operator $\hat{A}$ is the self-adjoint generator of $\hat{U}$. Clearly, $\hat{U}$ and $\hat{A}$ must commute with $\hat{H}$. Now suppose that the Hamiltonian $\hat{H}$ of a system is invariant under a set of symmetry operators $\hat{U}$. Then any of these symmetry operators acting on a ground-state (GS) wavefunction ψ_{ig} yields either the original GS wavefunction or another that is degenerate with it:

$$\hat{H}\psi_{ig} = E_g\psi_{ig}, \quad \hat{U}\hat{H}\psi_{ig} = E_g\hat{U}\psi_{ig}, \quad \hat{H}\hat{U}\psi_{ig} = E_g\hat{U}\psi_{ig} \tag{4.4}$$

At this point, it is important to recall two definitions from Group Theory.

- **Invariant Subspace:** Let $\mathscr{V}$ be a finite dimensional vector space with $dim(\mathscr{V}) \geq 1$, and let $\mathscr{T} \in \mathscr{L}(\mathscr{V},\mathscr{V})$ be an operator in $\mathscr{V}$. THen a subspace $\mathscr{U} \subset \mathscr{V}$ is called an invariant subspace of $\mathscr{V}$ if

$$\mathscr{T}u \in \mathscr{U} \qquad \forall u \in \mathscr{U} \tag{4.5}$$

- **Irreducible Representation:** An irreducible representation $(\mathscr{D},\mathscr{V})$ of a group G is a nonzero representation that has no proper nontrivial subprepresentation $(\mathscr{D}_u,\mathscr{U})$, with $\mathscr{U} \subset \mathscr{V}$ closed under the action of $\{\mathscr{D}(g) : g \in G\}$

Now, let the set $\{U_1,\hat{U}_2,...\}$ of symmetry operators and their inverses be closed under multiplication, forming a symmetry group. Any linear combination of degenerate eigenstates of $\hat{H}$ is another degenerate eigenstate, and the space of degenerate eigenstates is spanned by N_g orthonormal degenerate ground states that form the basis for a symmetry-invariant subspace of the state space, and the basis for an N_g-dimensional irreducible representation of the symmetry group. N_g is the degeneracy of E_g.

The statistical density operator for the ground state in the microcanonical (maximum entropy) ensemble at zero temperature is,

$$\frac{1}{N_g}\sum_{i=1}^{N_g} \left|\psi_{ig}\right\rangle\left\langle\psi_{ig}\right| \tag{4.6}$$

where $1/N_g$ is the probability of finding the system in the ith ground state. This is a product of a constant and the projection operator onto the subspace of degenerate ground-state wavefunctions, which is invariant when the same symmetry operator $\hat{U}$ is applied to all the ground-state wavefunctions. Starting from one ground-state wavefunction, all of those that are degenerate with it by symmetry (and not accidentally) can be generated by applying the symmetry operators to it. In this sense, "the symmetry of the ground state is the symmetry of the Hamiltonian", and the symmetry of the ground-state density is also the symmetry of the Hamiltonian. An important special case occurs when the ground-state is non-degenerate, as for typical closed-shell systems Then, in a stronger sense (i.e., for individual ground-state wavefunctions), the symmetry of the ground-state density is the symmetry of the Hamiltonian. While atoms that form chemical bonds are typically open-subshell, and their pure states may not have the spherical symmetry of the atomic Hamiltonian, the molecules that they form are typically closed-subshell (but not in every case).

The symmetry operators in $\{U_1, \hat{U}_2, ...\}$ may not all commute with each other, but those that do commute with each other can still be diagonalized along with $\hat{H}$. This is the typical case for open-shell non-relativistic (Coulomb interacting) atoms, where we can simultaneously diagonalize $\hat{H}, \hat{S}^2, \hat{S}_z, \hat{L}_z$. Since the energy of a real atom depends on the quantum numbers S and L that determine the eigenvalues $S(S+1)$ of the operator $\hat{S}^2$ and $L(L+1)$ of the operator $\hat{L}^2$, the true ground-state wavefunctions must also be eigenstates of the operators $\hat{S}^2$ and $\hat{L}^2$.

They can be but they do not have to be chosen as eigenstates of $\hat{S}_z$ and $\hat{L}_z$ for any choice of the z-axis. Thus the degeneracy of a multiplet is $(2S+1)(2L+1)$. For a given electronic configuration at the non-interacting level, the ground multiplet should have maximum possible S, and for that S maximum possible L (Hund's rule). The symmetry operators for rotation through angle about the z-axis in coordinate space are $e^{i\phi\hat{L}_z}$, and in spin space $e^{i\phi\hat{S}_z}$. The ground-state density will have the full symmetry of the Hamiltonian in thermal-equilibrium ground ensembles and non-degenerate pure ground states, but not necessarily for degenerate pure ground states.

Kohn-Sham ground-state spin-density functional theory in principle predicts the ground-state electron spin density and total energy of a real electronic system in the presence of a multiplicative and possibly spin-dependent external potential. The Kohn-Sham non-interacting wavefunction, often taken to be a single Slater determinant of spin orbitals that are eigenstates of the z component of an electron's spin, is intended to reproduce the spin densities (and non-interacting kinetic energy) of the interacting ground state, but it has been suggested that it need not otherwise be regarded as an approximation to a true ground-state wavefunction. From this perspective, it is hard to see why the Kohn-Sham

wavefunction should be constructed as an eigenstate of the operators $\hat{S}^2$ and $\hat{L}^2$. While "spin contamination" can be removed from a wavefunction by projection, that approach seems better justified in wavefunction theory than in DFT. For an atom, the Kohn-Sham non-interacting wavefunction [19] needs to be constructed as an eigenstate of the operators $\hat{S}^2$ and $\hat{L}^2$ only when that is required for the construction of an exact ground-state density. But that may in fact be required if we want a simple way to avoid symmetry-broken densities. It is known that singlet spin states of unbroken symmetry cannot have non-zero net spin densities, while symmetry-broken singlets have them. The work of Fertig and Kohn suggests [241] that eigenstates of $\hat{L}^2$ also have characteristic features in their densities.

Fertig and Kohn [241] argued that the total density of electrons in an atom with quantum number L can be expanded as a sum of spherical harmonic contributions with even ℓ in the range $0 \leq \ell \leq 2L$. That would be true both for the exact electron density and for a density constructed in a non-self-consistent Kohn-Sham approach using a spherically averaged Kohn-Sham potential. Although a self-consistent Kohn-Sham calculation with an approximate functional could bring spherical harmonic contributions with $\ell > 2L$ into the electron density, along with a non-spherical Kohn-Sham potential, they argued that those inappropriate contributions to the density would be small. They also showed that the exact Kohn-Sham potential of an atom is non-spherical, and has an expansion that includes non-zero spherical harmonics to infinite order. That the symmetry of the exact Kohn-Sham potential need not be that of the external potential seems to be consistent with Ref. [32]. Nevertheless, the most direct way to preserve the symmetry of the density is to constrain the approximate Kohn-Sham potential to have the symmetry of the external potential, and that is the approach we will follow here.

The exact density functionals constructed from a Levy constrained search over many-electron wavefunctions will have the exact degeneracies of the exact quantum mechanical problem (even in the Kohn-Sham scheme if all the densities of the degenerate ground states are non-interacting v-representable), but approximate functionals will not. In many cases, the approximate functionals will predict more accurate total energies for the densities of the broken-symmetry states. For example, the computationally efficient semilocal functionals best describe moderate correlation like that in the slowly-varying electron gas on which they are largely based. They can describe strong correlation in a symmetry-unbroken wavefunction only by converting it to moderate correlation in a symmetry-broken wavefunction, and density functional predictions of net spin densities can be re-interpreted as predictions of total density and on-top pair density [33]. How should we define the energy of an open-shell atom for the calculation of atomization energies of molecules and solids from approximate functionals? The standard approach is to choose the broken-symmetry atomic state of the lowest approximate energy. We see nothing wrong with

that, and it can lead to very accurate atomization energies when an accurate approximation like SCAN is used, as shown in Table IV with data from Ref. [36]. From the viewpoint of the density functional variational principle, this is the right thing to do, since it minimizes the approximated total energy functional via fully self-consistent calculations with possibly non-spherical Kohn-Sham effective potentials. (The predictiveness of SCAN for molecules and solids extends well beyond the atomization energies of Tables II and IV, but is not the main point of this article.)

Symmetry breaking in density functional theory often brings positive benefits, including more accurate energies from approximate functionals and physical insight into strong correlations that are only implicit in the symmetry-unbroken ground-state wavefunction but freeze out in the DFT total or spin densities [14, 231]. But symmetry breaking emerges with growing system size [14, 231], and is not expected to be important in small atoms. In the next section, we will show that spatial symmetry breaking in the first-row atoms occurs in approximate DFT, but that it can be energetically unimportant, and that the symmetries of the exact density can be preserved (not just restored by projection [30]) if that is needed.

Table 4.3: Mean absolute errors (MAEs) of the atomization energy for the six AE6 [34] sp-bonded molecules (SiH_4, SiO, S_2, C_2H_4, propyne, $C_2H_2O_2$ glyoxal, and C_4H_8 cyclobutane), in electron volts. The AE6 set was chosen [34] to be representative of the 109 atomization energies in Database/3 [35], which includes our O_2 and F_2 but not our B_2 and C_2. For the atoms and molecules, the self-consistent approximate Kohn-Sham wavefunction of lowest approximate energy is used, without imposing symmetries. Errors decrease from Hartree-Fock (numerical results from Ref. [34]) to DFT, and from the first to the third rungs of Jacob's ladder (numerical results from Ref. [36]) of approximations to the density functional for the exchange-correlation energy. These approximations are not fitted to any bonded system. A similar comparison for the formation energies of the 223 G3 molecules (including many molecules much larger than those in AE6) can be found in Ref. [18].

Hartree-Fock E_x	6.3
LSDA E_{xc}	3.3
PBE GGA E_{xc}	0.6
SCAN meta-GGA E_{xc}	0.1 ($\sim 1\%$)

4.4 How Energetically Important is Symmetry-Breaking in an Open-Shell Atom?

The symmetry of the many-electron Hamiltonian is the symmetry of the external potential, and the symmetry of the Kohn-Sham one-electron effective Hamiltonian is the symmetry of the Kohn-Sham effective potential. These symmetries are the same when the Kohn-Sham effective potential is constructed from an appropriately chosen equilibrium ensemble.

Following a path laid out by Ref. [233], we can preserve the spherical symmetry of the external potential (and thus of the full interacting Hamiltonian) for an atom in the Kohn-Sham potential (and effective Hamiltonian) by doing a self-consistent equilibrium-ensemble Kohn-Sham calculation with fractional occupation numbers. But the spherical densities do not belong to degenerate ground-state wavefunctions. The symmetry of the external potential dictates, but is not necessarily the same as, the possible symmetry of the wavefunction. So, we recover a symmetry-preserved density by replacing the fractional occupation numbers in the equilibrium ensemble by integers, as done in Ref. [233] and described in the first paragraph of section 4.2. That is only a first step toward symmetry preservation, but it is what we implement in our numerical calculations. The full process of symmetry preservation is presented in the next paragraph.

For atomic ground states, non-interacting many-electron eigenstates of $\hat{S}^2$ and $\hat{L}^2$ are not typically single Slater determinants. If we want to construct a single Slater determinant to represent a Hund's rule ground state in an atom, we need to focus on the states with $\hat{S}^2$ and $\hat{L}^2$ eigenvalues $M_S = \pm S$ and $M_L = \pm L$, since these have single fully detailed electronic configurations. The spin orbitals then must be constructed as complex, current-carrying products of radial functions and spherical harmonics (complex linear combinations of the first-step symmetry-preserving orbitals). Since the densities of the spherical harmonic orbitals have cylindrical symmetry about the z axis, the total density also has this symmetry (as it has in first-step symmetry preservation). The presence of a non-zero current density implies that meta-GGAs like SCAN must be implemented in a way that is not standard in most electronic structure codes. We defer these later steps to future work.

We start with a self-consistent calculation of the spherical, spin-polarized equilibrium ensemble density. For example, we occupy the carbon atom with $p_x^{2/3} p_y^{2/3} p_z^{2/3}$ electrons in S_α and no electrons in S_β. Then we find the self-consistent Kohn-Sham potential that this density produces, which is also spherical, like the external potential of an atom. We take the atomic orbitals for that potential, then occupy them with integer occupation numbers. Finally, we compute the corresponding energy for these orbital occupations, for comparison with the energy of the self-consistent broken-symmetry solution. This procedure has been carried out for the oxygen atom with $p_x^1 p_y^1 p_z^1$ electrons in S_α and

$p_x^{1/3} p_y^{1/3} p_z^{1/3}$ electrons in S_β, the boron atom with $p_x^{1/3} p_y^{1/3} p_z^{1/3}$ electrons in S_α and no electrons in S_β, the fluorine atom with $p_x^1 p_y^1 p_z^1$ electrons in S_α and $p_x^{2/3} p_y^{2/3} p_z^{2/3}$ electrons in S_β. Table 4.4 compares the first-step symmetry preserved and -broken energies of the atoms under study.

Table 4.4: Effect of first-step non-self-consistent spatial symmetry preservation on the atomic energies of B, C, O, and F with the use of three density functionals SCAN, PBE and LSDA. All densities are non-spherical. Sym-Br stands for symmetry-broken or self-consistent and Sym-P1 for first-step non-self-consistent spatial symmetry- preserved computations. Energies are in Hartree, unless otherwise specified.

Atom	**SCAN**	**PBE**	**LSDA**
B (Sym-P1)	-24.6364	-24.6085	-24.3520
B (Sym-Br)	-24.6393	-24.6100	-24.3528
Diff (eV)	0.08	0.04	0.02
C (Sym-P1)	-37.8343	-37.7939	-37.4644
C (Sym-Br)	-37.8371	-37.7953	-37.4653
Diff (eV)	0.08	0.04	0.02
O (Sym-P1)	-75.0620	-75.0041	-74.5188
O (Sym-Br)	-75.0663	-75.0071	-74.5210
Diff (eV)	0.12	0.08	0.06
F (Sym-P1)	-99.7328	-99.6613	-99.0979
F (Sym-Br)	-99.7371	-99.6644	-99.1003
Diff (eV)	0.11	0.08	0.06

The result that emerges from the data is that the energy difference between first-step symmetry-preserved and symmetry-broken densities is small. Table 4.4 highlights that spatial symmetry breaking lowers the energy of an atom, as expected, but only inconsequentially, usually much less than the errors of approximate DFT atomization energies. The overall smallness of the energy differences quantifies the conclusions of Fertig and Kohn [241]. Thus, the standard way of calculating atomization energies from Kohn-Sham theory, employing a single Slater determinant with self-consistent symmetry breaking to minimize the energy, is supported by this investigation. Table 4.4 also shows that the energy differences grow from LSDA to PBE to SCAN, in keeping with the increasing density-sensitivity of the functionals as already noted in the discussion of Table 4.1.

Next, we consider symmetry breaking in both space and spin. To make the Kohn-Sham effective potential spin-independent, like the external potential, we must begin with a spin-unpolarized equilibrium ensemble. We occupy the carbon atom with $p_x^{1/3} p_y^{1/3} p_z^{1/3}$

electrons in S_α and $p_x^{1/3} p_y^{1/3} p_z^{1/3}$ electrons in S_β. Then we take the atomic orbitals for that potential and occupy them with integer occupation numbers for the correctly spin-polarized atom. Finally, we compute the corresponding energy for these orbital occupations, for comparison with the energy of the self-consistent broken-symmetry solution. This procedure has been carried out for the oxygen atom with $p_x^{2/3} p_y^{2/3} p_z^{2/3}$ electrons in both S_α and S_β, the boron atom with $p_x^{1/6} p_y^{1/6} p_z^{1/6}$ electrons in S_α and S_β, and the fluorine atom with $p_x^{5/6} p_y^{5/6} p_z^{5/6}$ electrons in S_α and S_β. Table 4.5 compares the symmetry- preserved and -broken energies of the atoms under study.

Table 4.5: Effect of first-step non-self-consistent space-spin symmetry-preserved atomic energies of B, C, O, and F atoms with the use of three density functionals SCAN, PBE and LSDA. All densities are non-spherical. Sym-Br stands for symmetry-broken or self-consistent and Sym-P1' for first-step space-spin symmetry-preserved computations. Note that the Sym-Br or self-consistent total energies are the same as in Table 4.4, but the Sym-P1' total energies are higher than the Sym-P energies because more symmetries are preserved in Sym-P. Energies are in Hartree, unless otherwise specified.

Atom	**SCAN**	**PBE**	**LSDA**
B (Sym-P1')	-24.6358	-24.6083	-24.3518
B (Sym-Br)	-24.6393	-24.6100	-24.3528
Diff (eV)	0.10	0.05	0.03
C (Sym-P1')	-37.8315	-37.7927	-37.4630
C (Sym-Br)	-37.8371	-37.7953	-37.4653
Diff (eV)	0.15	0.07	0.06
O (Sym-P1')	-75.0585	-75.0016	-74.5157
O (Sym-Br)	-75.0663	-75.0071	-74.5210
Diff (eV)	0.21	0.15	0.14
F (Sym-P1')	-99.7319	-99.6605	-99.0969
F (Sym-Br)	-99.7371	-99.6644	-99.1003
Diff (eV)	0.14	0.11	0.09

The averaged difference between these two approaches in Table 4.5 is 0.15 eV for SCAN, 0.09 eV for PBE, and 0.08 eV for LSDA. Relative to the energy of our self-consistent calculation, spin-symmetry preservation raises the energies of our four atoms and would also raise the energies of our two triplet molecules. A standard way to preserve spin symmetry is the restricted open-shell formalism, as described for Hartree-Fock theory in Ref. [40]. It employs the same up- and down-spin orbitals for spin-paired electrons in an open-shell configuration, like the approach in our Table 4.4. The approach of Table 4.4

restores spin symmetries in a simple non-self-consistent way, and also makes a first step toward restoring spatial symmetries.

CHAPTER 5

INCORPORATION OF DENSITY SCALING CONSTRAINT IN DENSITY FUNCTIONAL DESIGN VIA CONTRASTIVE REPRESENTATION LEARNING

5.1 Introduction

Density functional theory (DFT) is an indispensable tool in computational chemistry and materials science due to its combination of efficiency and accuracy [14, 19]. As the standard computational method that is widely applied in physics, chemistry, and materials research, DFT has achieved high prediction accuracy enabled by the continued development of approximations of the exchange-correlation (XC) energy as a functional of electron density [35, 40]. An appropriately approximated density functional enables more accurate first-principles calculations for molecules and material systems on a larger scale. In different forms of approximations, the XC functionals must satisfy several exact conditions and constraints, such as coordinate scaling [21], spin scaling [22] and derivative discontinuity [242]. So far, all popular approximations suffer from systematic errors that arise from the violation of mathematical properties of the exact functional. It is expected that the performance and generality of density functionals can be improved by satisfying these constraints. For instance, the recently developed strongly constrained and appropriately normed (SCAN) functional [40] that satisfied 17 exact constraints achieved great performance for both molecules and solids.

There has been a growing interest in applying machine-learning (ML) in physics, chemistry, and material science, with the aim of achieving the same or even higher prediction accuracy for molecules and materials with much less computational cost compared to first principles simulations. Recently, ML has been applied to parametrize XC functionals without domain knowledge of humans by using various methods such as kernel ridge regression (KRR) [243], fully connected neural networks (NN) [244, 245] and convolutional neural networks (CNN) [246]. Being trained in a supervised manner, these ML models are highly accurate across a small set of molecule systems similar to those on which the models are trained, while in many cases they show a worse performance on larger molecular datasets as they do on small ones. Neither of them demonstrate the same level of universality compared to conventional XC functionals.

There has been plenty of effort being devoted to leveraging physical constraints in ML of XC functionals. In a previous work by Lei et al. [246], by using CNN as encoders, rotational invariant descriptors were extracted and projected on a basis using spherical harmonic kernels. In another work by Hollingsworth et al. [247], it was found that the scaling property [21], which is one of the exact conditions that the exchange energy must satisfy, can be utilized to improve the machine learning of XC functionals. The study is limited to one-dimensional systems which lacks the generalizability to two- and three-dimensional systems. More recently, another exact condition - derivative discontinuity - was incorporate into the NN-based XC functional design [248], while the study is again limited to one-dimensional systems. A more recent work has demonstrated that the fundamental limitation can be overcome by training a neural network on molecular data and on fictitious systems with fractional charge and spin [75], and the resulting NN-based functional DeepMind 21 [75] demonstrated the universality and greatly improved predictive power for molecule energetics and dynamics. These works use data augmentation to improve model performance by directly increasing the amount of labeled data.

However, increasing the amount of data is not always possible due to the computational cost. Going beyond data augmentation, self-supervised learning has gained popularity because of its ability to avoid the cost of annotating large-scale datasets. It adopts self-defined pseudo labels as supervision and uses the learned representations for downstream tasks. Self-supervised learning has been widely used in image representation learning [249], natural language processing [250], and applied in molecular machine learning [251, 252]. Specifically, contrastive learning (CL) has recently become a dominant branch in self-supervised learning methods for computer vision, natural language processing, and other domains [253]. It aims at embedding augmented versions of the same sample close to each other while trying to push away embeddings from different samples in the representation space. The goal of contrastive representation learning is to learn such an embedding

space in which similar sample pairs stay close to each other while dissimilar ones are far apart, and the CL process can be applied in both unsupervised and supervised settings [254]. In this work, we will explore the incorporation of physical constraints in density functional learning through contrastive learning.

One of the most important and fundamental constraints for the exchange energy is derived from the principle of uniform scaling [21]. Consider an electron density distribution $n(\mathbf{r})$ and a scaled density $n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r})$ (where γ is a scaling factor), several important exact constraints on exchange and correlation energy functionals can be written using the scaled density. In this work, we focus on the exchange energy $E_x[n]$, and the scaling property must satisfy the following requirements: $E_x[n_\gamma] = \gamma E_x[n]$. This important constraint is satisfied exactly in almost all human-designed density functionals, whether non-empirical or semi-empirical. As a chemical example, atomic one-electron ions of nuclear charge Z are scaled versions of the hydrogen atom with scale factor $\gamma = Z$. The exchange energy, $-5Ze^2/16a_0$, in this case cancels the Hartree electrostatic interaction of the density with itself. Using this constraint as an outstanding example, in this work we propose a schematic approach to incorporate any physical constraints (represented by equalities) via contrastive representation learning into the NN-based model design.

Specifically, we pre-trained an electron density encoder by maximizing the similarity between molecular electron density and its scaled version with a randomly chosen scaling factor, within the framework of SimCLR [255], which is a widely used framework for contrastive learning of image pretraining. To obtain an encoder that gives similar representations (while different by a scaling factor) for scaled and unscaled electron densities, we added a scaling factor predictor component to the framework. The pre-trained encoder was then transferred to the downstream task to predict the exchange energies from electron densities. It is shown that the model pretrained contrastively predicts exchange energies that satisfy the scaling relation, while the model trained on unaugmented dataset gives poor predictions. We will show that contrastively learned encoders are capable of encoding molecular electron density with less labeling cost from the fact that they give comparable predictions by fine-tuning using only a small percentage of labeled data, compared to the model trained on the whole labeled dataset by supervised learning. This shows that contrastive learning using constraints can enhance the understanding of DFT theory for neural network models with a small amount of labeled data while generalizing the application of NN XC functionals in a wide range of scenarios which are not always available experimentally but theoretically justified.

5.2 Results

5.2.1 Electron Density Encoder

In this work, the Residual Network (ResNet) was used as the electron density encoder, which is the most commonly used network in image recognition. With the deeper and deeper the neural networks, effective learning becomes more challenging due to the gradient vanishing or exploding problem [256, 257], which makes traditional models using convolutional neural network layers reach a limit of performance when the number of layers increases. In 2016, He et al. proposed using skip-connection that allows direct connection from the input layer to the output. By skipping intermediate layers, if there is a gradient issue for these layers, the model is still able to learn the identity map, which ensures that the gradient is at least equal to 1. Instead of learning the mapping H between input x and target y, residual networks aim to learn the residual F:

$$F(x) := H(x) - x \tag{5.1}$$

In the worst case, a trivial result is learned such that $F(x) = 0$, the mapping H is the identity mapping $H(x) = x$. Such kind of architecture enables the learning ability of neural networks that are extremely deep.

5.2.2 Contrastive Learning of Uniform Density Scaling Property

Contrastive learning (CL) is a self-supervised learning (SSL) strategy that learns useful representations using unlabeled data, by manually designing pre-training tasks with automatically generated labels. Typically, when applied in image recognition, data augmentations such as random shifting, random cropping and random rotation are applied to generate different views of images. The raw and augmented images are then passed to an image encoder to generate hidden representations that are passed to a projection head projecting representations onto a high dimensional unit sphere. The projected representations are used to calculate contrastive loss that maximizes the similarity between projected representations of the same input image, while minimizing the similarity between those of different images. By minimizing contrastive loss and updating the model parameters through backpropagation, the image encoder is aware of that the different views are from the same raw image, which introduces invariance to the model for imperfect inputs. Intuitively, encoder trained by contrastive learning groups different views of the same image into the same cluster while pushing clusters from different images far away from each other.

In this work, we intend to design a pre-training task such that the electron density encoder is aware of the uniform density scaling property. In order to do so, unscaled and

scaled electron density on a fixed-size spatial grid are generated using PySCF code with low computation cost, represented as three-dimensional arrays x_i, $\tilde{x}_{i\gamma} \in \mathbb{R}^{(d\times d\times d)}$, where the scaling factor γ is randomly chosen from four different scales: $1/3, 1/2, 2, 3$. Electron density arrays are encoded as hidden representations $h_i = f(x_i)$, $\tilde{h}_{i\gamma} = f(\tilde{x}_{i\gamma}) \in \mathbb{R}^m$ through the density encoder that is a map $f : \mathbb{R}^{(d\times d\times d)} \to \mathbb{R}^m$ to be learned. The hidden representations are then projected as a set of points $z_i = g(h_i) \in \mathbb{R}^n$ on a high dimensional unit sphere by a map $g : \mathbb{R}^m \to \mathbb{R}^n$ $(n < m)$ that is a multilayer perceptron (MLP). For a batch of N molecules, output $Z \in \mathbb{R}^{(2N\times m)}$ contains projected representations of unscaled and unscaled densities. The loss is then calculated by the normalized temperature-scaled cross entropy (NT-Xent) loss [255] that is defined as:

$$l_{i,j} = -\log \frac{\exp\{(z_i \cdot z_j)/\tau\}}{\sum_{k=1,\, k\neq i}^{2N} \exp\{(z_i \cdot z_k)/\tau\}} \tag{5.2}$$

where temperature factor τ is a positive real number, and the exponential term when $k = i$ is excluded in the summation in the denominator to ensure that the loss is zero if dissimilar projected representations are perpendicular.

Indeed, for $z_i \cdot z_j \neq 0$, $z_i \cdot z_k = 0$ $(k \neq j)$,

$$l_{i,j} = \log\left(1 + \frac{\sum_{k=1,\, k\neq i}^{2N} \exp\{(z_i \cdot z_k)/\tau\}}{\exp\{(z_i \cdot z_j)/\tau\}}\right) = 0 \tag{5.3}$$

For a batch of N molecules, z_{2k-1} and z_{2k} are corresponding projected representations of unscaled and scaled densities of the same molecule. Notice that the loss function is asymmetric: $l_{ij} \neq l_{ji}$, the total loss is

$$L = \frac{1}{2N} \sum_{k=1}^{N} (l_{2k-1,\, 2k} + l_{2k,\, 2k-1}) \tag{5.4}$$

The loss is zero when projected representations of different molecules are perpendicular to each other, which ensures that dissimilar samples are pushed far apart from each other.

In the original SimCLR framework [255], augmented and unaugmented view of the same input form positive pairs, while those of different inputs form negative pairs. We need to point out that without any modules added to distinguish positive pairs, the encoder trained would be too "lazy" to learn different representations for the two "views" of the same input. Because the simplest map f that minimize the loss satisfies $f(\tilde{x}_{i\gamma}) = f(x_i)$. Hence, a module predicting the scaling factor from two hidden representations of the same molecule is added. The final loss of the contrastive pretraining task is the summation of these two losses. The workflow of pretraining task is shown in Fig. 5.2.

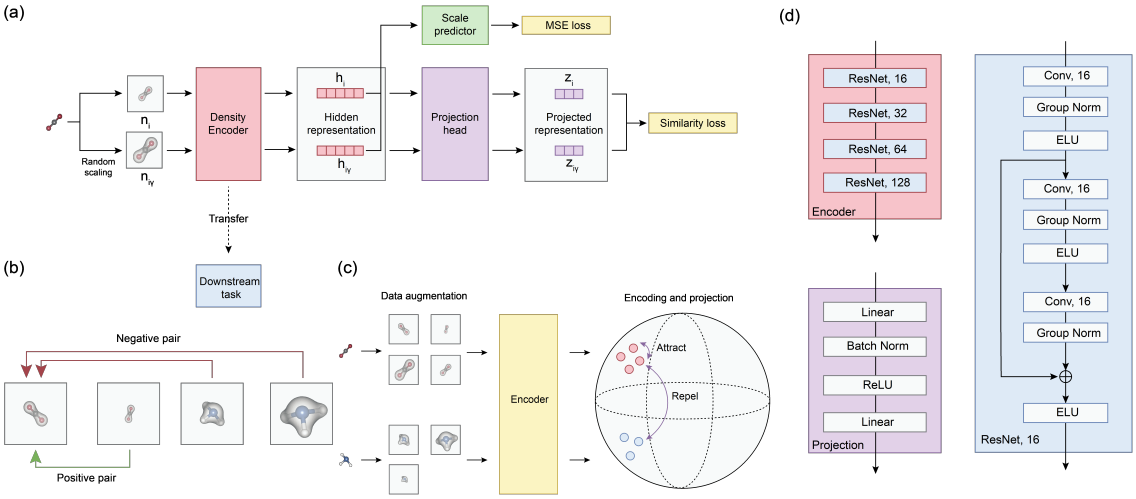

Figure 5.1: (a) The workflow of this work. For arbitrary molecule, an unscaled and a scaled electron density are fed into the density encoder to obtain hidden representations. The subsequent modules are divided in to two parts: a projection head that produces projected representations, from which the contrastive similarity loss is calculate; a scale predictor that predicts scaling factor from the hidden representation pairs, from which the mean squared error loss is calculate. (b) The two electron densities from the same molecule form positive pairs, while those from different molecules form negative pairs. (c) The visualization of general contrastive learning. Multiple "views" of the same input molecule are generated by data augmentation. After encoding and projection, representations from the same molecule attract each other, while those from different molecules repel each other. (d) The architecture of density encoder, projection and ResNet building block.

The cosine similarity of learned projected representations z and $\tilde{z}$ for a batch of 32 molecules are shown in Fig. 5.2 (a). As expected, the cosine similarity has maximum value for positive pairs – unscaled and scaled densities of the same molecules, while having close to zero value for negative pairs – densities of different molecules. We further verify that projected representations are well separated from each other by computing the t-distributed neighbor embedding (tSNE). In Fig. 5.2 (c), two examples of molecules, learned projected representations and predictions on scaling factors are shown. The best model achieves 0.01976 contrastive loss and 2×10^{-4} MSE for scaling prediction.

5.2.3 Comparison of Performance of Supervised Learning and Contrastive Learning

Supervised learning of neural networks is one of the most widely used machine learning strategies in material science. By training with a large amount of labeled data, which

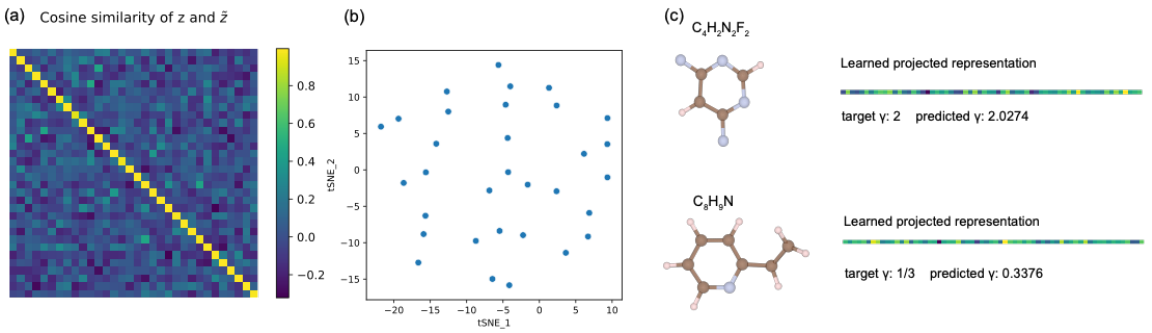

Figure 5.2: (a)The cosine similarity between learned projected representation of unscaled and scaled densities for a batch of 32 molecules. Each element in the matrix is computed as $\cos\left(z_i \cdot \tilde{z}_j\right) := z_i \cdot \tilde{z}_j$. The brighter it is, the closer the value is to 1. (b) The t-distributed stochastic neighbor embedding (tSNE) of 32 learned projected representations. (c) Two examples of molecules, the corresponding learned projected representations, and predictions on scaling factors.

means that each input has the corresponding target, the model can give predictions with a small discrepancy with the true targets. However, good performance on a dataset does not guarantee equally good performance on other datasets. In this section, we will show that the model trained by supervised learning on unscaled density data only may achieve a very low prediction error for predicting exchange energies from molecular electron densities but at the same time demonstrates a large prediction error for scaled densities. This clearly shows that the model trained on unscaled density dataset with supervised learning does not understand the uniform scaling property that exchange energy functionals must satisfy.

Using the data-driven approach, the mapping of molecular electron density to the exchange energy, is directly learned in a supervised manner by feeding electron densities, with the corresponding exchange energies calculated from first-principles calculations as labels, to an electron density encoder. Electron density in three-dimensional space is represented by a three-dimensional array, with the dimension along each axis equals the grid dimension along the same axis. Encoding and decoding of volumetric data in three-dimensional space has been previously studied in 3D-UNet [258], with a DoubleConv layer consisting of two subsequent 3D convolutional layers as the building block. In the same 3D-UNet framework, instead of DoubleConv, residual networks can be used as the building block to extract useful information from raw three-dimensional volumetric data [259]. In this work, the mapping of electron density to the exchange energy will be learned, so only the encoder part will be adopted. The encoder consists of several connected building block layers, being either DoubleConv or ResNet (see Methods).

The architecture of the encoder is shown in Fig. 5.1 (d). A hidden representation

that captures density-energy correlation is learned and fed to a subsequent fully connected prediction layer to give a single value prediction on the exchange energy. The electron densities of molecules with a scaling factor equal to one are included in the dataset. For reliable evaluation of the models, the dataset is split into 80% and 20% as training and validating datasets. The training set is employed to train the model for 500 epochs by minimizing the mean squared error (MSE) loss, and the model is then applied to validate the performance on validate set using the mean absolute error (MAE) as the measure.

To investigate whether the model trained with only unscaled densities understand the uniform density scaling property, we test its performance on both unscaled and scaled density dataset. As shown in Fig. 5.3 (a), the relative difference between predictions and targets on unscaled dataset is as small as 0.45 eV. But the model does not provide reliable predictions for the scaled dataset. This indicates that the models trained in a supervised manner in general literatures may give wrong predictions for scaled density that do not satisfy the uniform density scaling property, although they may achieve high accuracy on unscaled density dataset. This motivates us to apply contrastive learning in a pretraining task to give the model understandability of scaling property.

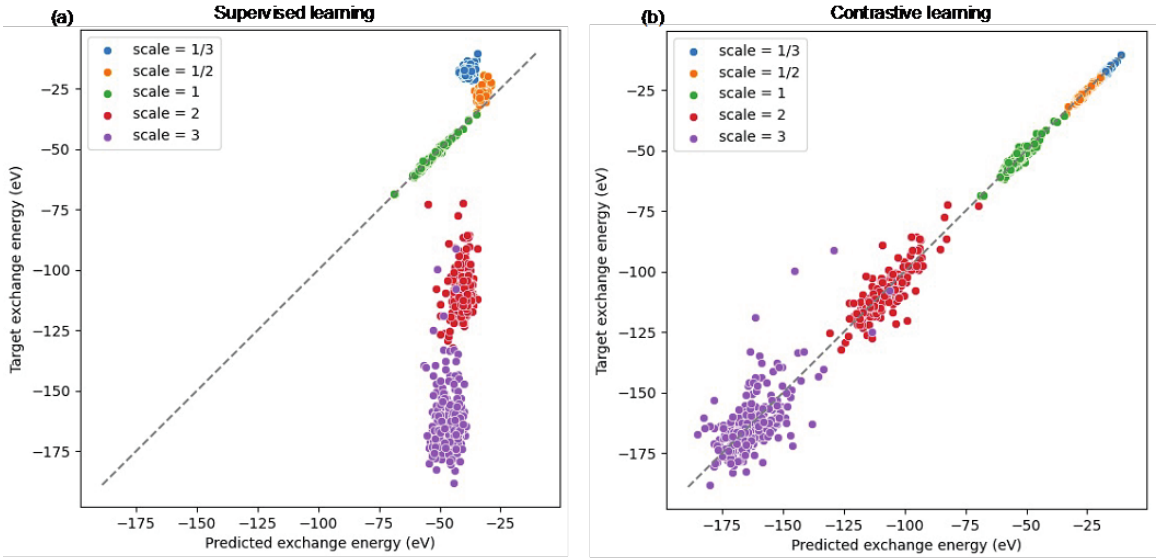

Figure 5.3: Comparison of performance between supervised learning model and contrastive learning model on datasets with different scales. (a) Model trained by supervised learning on unscaled dataset (green) gives poor prediction on other scaled datasets. (b) Model trained by contrastive learning give relatively good predictions on all datasets.

5.2.4 Contrastive Learning Model Performance with Different Label Percentages

Now we investigate the model for predicting exchange energy from electron density. The density encoder part of the model is transferred from contrastive pretraining task. The model is also trained from scratch and the performance of which is compared to the transferred model. When finetuning the transferred model, we adopted training sets with 80%, 60%, 40% and 20% labeled data. Surprisingly, the transferred model outperforms the model trained from scratch with label percentage as low as 60%. This shows that our methods can reduce the need of large amount of data while still have better performance. More than that, model trained with our method enable the understandability of the uniform density scaling property. As shown in Fig. 5.4, predicted and target exchange energy still have strong linear correlation as label percentage decreases. Note that for the case of 20% label percentage, the model uses the same amount of labels as that of the supervised learning task in previous section. But the dramatic difference of performance between models shown in Fig. 5.3 shows the understandability of uniform scaling property of models trained by our method.

5.3 Discussion

In this work [260], contrastive learning is adopted to pretrained electron density encoder to incorporate the uniform density scaling property for exchange energy predictions. Generated from first-principles calculations, the scaled and unscaled electron densities of molecules from QM9 dataset are used to contrastively trained the electron density encoder. Scaled and unscaled densities of the same molecule are recognized as similar, while those from different molecules are recognized as dissimilar. The pretrained model achieves 0.01976 contrastive loss. It also predicts scaling factors from hidden representations of scaled and unscaled densities, with 2×10^{-4} MSE accuracy. The encoder is then transferred to a downstream task to predict DFT calculated exchange energies from electron densities with different scaling parameters. With contrastive learning as pretraining method, our model predicts exchange energies that satisfy the uniform scaling property, while the model trained using unscaled densities in a supervised manner give bad predictions for scaled densities. This shows that contrastive learning helps model learn physical principle.

Table 5.1: The MAE of model with ResNet and DoubleConv as density encoder for predicting exchange energy. Performance is tested from scratch, with 80%, 60%, 40% and 20% labeled data.

Label Percentage	ResNet (16, 32, 64, 128)	DoubleConv (32, 64, 128)
Supervised Learning		
100%	2.2673	2.6107
Contrastive + Transfer Learning		
80%	1.8265	2.2000
60%	2.0353	2.2319
40%	2.3249	2.5910
20%	2.5472	2.9377

5.4 Methods

Molecular electron density dataset We chose 10k molecules from the QM9 dataset [261] by imposing the following criteria:

- each molecule contains less than 20 atoms,
- each molecule does not contain atoms with an atomic number larger than 36 (element Kr),
- the size of each molecule is less than 12 angstroms,
- the DFT calculated exchange energy of the molecule should be greater than -200 eV.

Molecular density matrices are calculated by DFT with PBE functional 3 as implemented in the PySCF package [262]. To prepare grid-like input data with fixed dimensions, we project the density matrices onto real space grid points with shape (65, 65, 65) on a fixed size cube centered at the origin with a length of 40 angstroms. The number of grid points is set to odd integers to include the origin. A larger grid with shape (129, 129, 129) is also used to construct more detailed density data. But due to the limit of storage for the whole dataset, an average pooling down-sampling pre-process is applied to reduce the grid dimensions from 129 to 65. A comparison of the results using these two grids is given in later sections. The projection of density matrices on grids in three-dimensional space is done by using PySCF code. The exchange energies are calculated using the same setting as the NWChem code [236].

5.4.1 Training and Evaluation of Supervised Learning Task

To find out the best model that encodes the electron density, two different types of building block layers: ResNet and DoubleConv, were used to build the density encoder. The model was built and trained using Pytorch-Lightning package [263] which is a framework based on Pytorch package [1]. The whole dataset is split into 80% and 20% for training and validating, respectively. Training loss is backpropagated to update the model parameters by Adam optimizer [264] with learning rate 0.001. The best model was chosen to be that with smallest MAE after 500 epochs [260].

5.4.2 Training and Evaluation of Contrastive Learning Task

The dataset consists of electron densities of 10k molecules chosen from QM9 dataset [261]. Each raw electron density is augmented by a scaled one with the scaling factor randomly chosen from $1/3, 1/2, 2, 3$. As a result of hyperparameter searching, ResNet with feature maps (16, 32, 64, 128) and DoubleConv with feature maps (32, 64, 128) are chosen for comparison of performance on downstream task. The whole dataset is split into 80% and 20% for training and validating, respectively. The total training loss is the summation of contrastive loss and scaling factor predicting loss, which is then backpropagated to update the model parameters by Adam optimizer [264] with learning rate 0.001. The best model was chosen to be that with smallest total loss after 1000 epochs.

5.4.3 Training and Evaluation of Downstream Task

The dataset consists of raw and four augmented electron densities for four scales: $1/3, 1/2, 2, 3$ of 10k molecules chosen from QM9 dataset, resulting in a dataset containing 50k electron densities.

5.5 Data Availability

The code and data for this work can be found at Github of [260].

5.5.1 Data Availability Statement

The authors declare that the main data supporting the findings of this study are available within the paper and its Supplementary files. Other relevant data are available from the

[1] A. Paszke, et al., in Proceedings of the 33rd International Conference on Neural Information Processing Systems (Curran Associates Inc., 2019), pp. Article 721.

corresponding authors of Ref. [260] upon reasonable request.

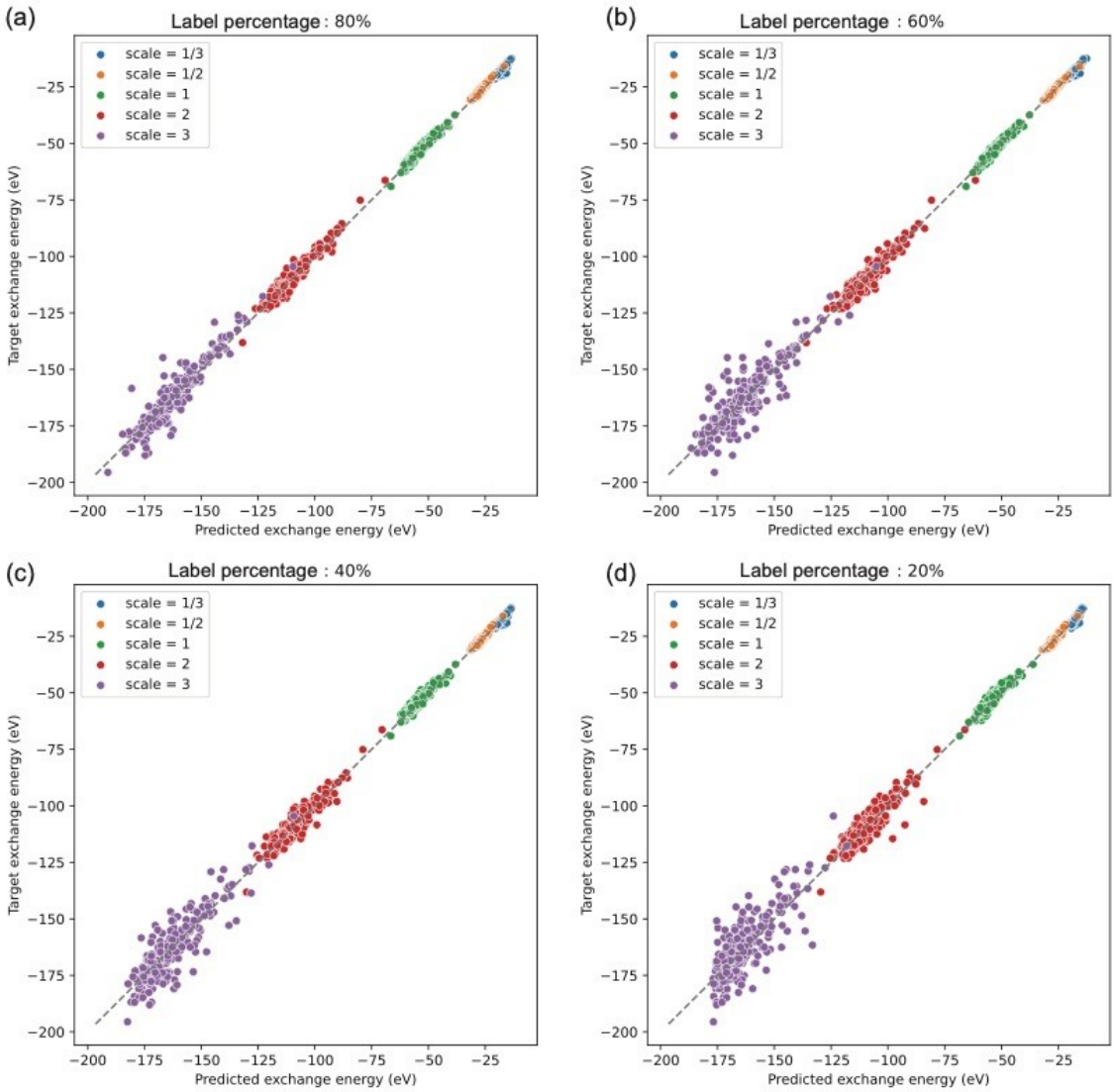

Figure 5.4: Prediction vs. target for four model finetuned with four different label percentages. As label percentage decreases, the model is still able to give relatively good predictions.

CHAPTER 6

EPILOGUE

In the first chapter, we show that the complicated interconnection between the electronic and nuclear motions of a many-body system can be simplified through some approximations. We then reformulate the electronic motion problem in the framework of density functional theory and discuss in brief, the features of various rungs of density functional approximations.

In the second chapter, we have extended our vdW correction model, vdW-dZK, to technologically important material systems: graphene adsorbed on monolayer and multilayer substrates of layered materials, such as G/MoS_2, G/WS_2, G/$MoSe_2$, G/WSe_2, G/graphite and G/h-BN-bulk, by including the C_4 terms and multi electrostatic image effects due to the two surfaces of the substrate slabs in the vdW interaction calculations. In these systems, without long-range vdW corrections, the PBE GGA strongly underestimates the binding and strongly overestimates the equilibrium distance, while the SCAN meta-GGA also underestimates the binding (but less severely) and is rather accurate for the equilibrium distance.

For the system of G/graphite, both the dZK-corrected SCAN and PBE methods give very accurate binding energies, very close to the RPA values. The predicted equilibrium distances from the dZK-corrected SCAN are also very accurate, while those predicted by the dZK-corrected PBE are about $0.2-0.3\,\AA$ larger than the RPA value. SCAN+rVV10 overestimates the binding energy for the G/graphite system. but gives almost the same equilibrium distance as that of the RPA. Since SCAN+rVV10 usually overbinds, the predicted binding energies from SCAN-vdW-dZK and PBE-vdW-dZK may be closer to the true values.

SCAN-vdW-dZK and PBE-vdW-dZK give approximately the same results for the monolayer and multilayer system considered here. The predicted binding energies from the two dZK-corrected methods are roughly 20-30% lower than those from the SCAN+rVV10, and 10-20% lower than those from PBE+rVV10. The biggest differences come from

graphene adsorbed on monolayer systems. While both dZK-corrected SCAN and PBE predict relatively large differences in the binding energies between monolayer and multilayer systems, the rVV10-corrected ones give relatively smaller differences among those systems. The C_4-term contributions can be 40% of the total vdW interactions and the C_5 term contributes about 10%. The effects of images and the back surfaces of slabs can contribute about 4-10%. The vdW interaction energy power law exponents from the vdW-dZK model for graphene adsorbed on MoS_2 multilayer show slowly varying decay to the pairwise -4 with increasing large D, very similar to those obtained from RPA and RG (renormalization group) approaches by Ambrosetti et al. [168]. Both PBE-vdW-dZK and SCAN-vdW-dZK give a greater increase in interlayer binding energy when the TMD substrate changes from monolayer to 4-layer for the graphene/TMD adsorption systems. This is consistent with the relevant EELS experimental results of increased dielectric response from substrates with increasing substrate layer numbers. rVV10 and vdW-dZK are both long-range van der Waals corrections to semilocal density functional approximations. rVV10 requires only a knowledge of the electron density, but it is limited to pairwise interactions between volume elements of that density. vdW-dZK is more complicated and requires material-specific input parameters, but seems to account better for screening and for non-pairwise interaction.

In the third chapter, we have developed a vdW interaction model for molecules adsorbed on curved cylindrical conducting surfaces, and have combined this model with the semilocal density functionals PBE and SCAN. The resulting PBE-vdW-dZK and SCAN-vdW-dZK are applied to NH_3 and NO_2 adsorbed on carbon nanotubes. The results from PBE-vdW-dZK and SCAN-vdW-dZK are also compared with those from vdW nonlocal functionals, such as SCAN+rVV10 and PBE+rVV10. The PBE functional captures almost no vdW interaction and underestimates the binding energies for the two systems. Even without the inclusion of long-range vdW interactions, SCAN can capture intermediate vdW interactions and gives much improved binding energies. When combined with the vdW-dZK model and vdW functional rVV10, both PBE and SCAN improve their descriptions for these systems in which vdW interaction is important.

Generally speaking, the binding energies from PBE+rVV10, SCAN+rVV10, PBE-vdW-dZK, and SCAN-vdW-dZK are roughly the same, about 70-115meV for CNT+NH_3 and 300-490meV for CNT+NO_2, respectively. The results from PBE-vdW-dZK and SCAN-vdW-dZK are closer to each other than are those of PBE+rVV10 and SCAN+rVV10. For CNT+NH_3, PBE-vdW-dZK and SCAN-vdW-dZK give binding energies of about 103-114 meV. For CNT+NO_2, PBE-vdW-dZK and SCAN-vdW-dZK give binding energies of about 422-490 meV. Comparatively, for CNT+NH_3, PBE+rVV10 and SCAN+rVV10 give binding energies of about 70-80meV, while for CNT+NO_2, they give about 300-420meV. The relatively closer results from PBE-vdW-dZK and SCAN-vdW-dZK indicate the consis-

tency of our developed vdW-dZK model for curved surfaces. Due to the relatively large systems ($>$ 80 atoms) considered here, there is no high-level computational result. The available reference data from the literature can just serve as a rough guideline for the binding energies of the systems, and more reliable reference values would yield a more reliable damping factor $\bar{b}$ in Eq. 3.11. However, our vdW-dZK combined with PBE and SCAN gives results closer to this guideline than other methods considered here. It seems that the results from PBE-vdW-dZK and SCAN-vdW-dZK could be more realistic. All methods, including PBE, SCAN, PBE+rVV10, SCAN+rVV10, PBE-vdW-dZK, and SCAN-vdW-dZK, give approximately the same binding energies for the two adsorption configurations (types I and II) for the two systems. This may imply that the two adsorption sites have approximately the same adsorption stability. The exponent of the vdW interaction power law from our vdW-dZK model for CNT+NH_3 and CNT+NO_2 systems is about 0 at short distance, largely due to the damping factor. At distances D about $20-50\,\AA$, the exponent is about -4.0 to -4.5, and not the pairwise limit -5. This feature is very similar to the one calculated by RPA and RG (renormalization group) approaches. This similarity provides further support to our methods.

In the fourth chapter, we show that approximate non-empirical density functionals become not only more accurate for total energies and their differences but also more sensitive to the density as we go from LSDA to PBE to SCAN. In this sequence, the functionals better approximate the exact constrained search over correlated wavefunctions, and not the exact constrained search over ensembles. (For open systems of fluctuating electron number, which can be described only by ensembles, the large errors made by such functionals for total energies and their differences have long been known) For accurate atomization energies of molecules from these functionals, and especially from SCAN, the densities of the open-shell atoms should not be sphericalized (ensemble-averaged over degenerate states). We have also found (Fig. 4.1) that the improvement in self-consistently-calculated atomization energy from LSDA to PBE is mainly an overall reduction of overbinding, while that from PBE to SCAN is more properly system specific for the normally-correlated molecules studied here.

The work of Fertig and Kohn suggests that, to yield features of the exact ground-state density, an approximate Kohn-Sham non-interacting wavefunction of an atom can be constructed as an eigenstate of the square of the total angular momentum operator, just as the true or interacting wavefunction is. The spatial symmetry of the interacting ground-state wavefunction can be preserved from fractional to integer occupation numbers. Full preservation of spatial symmetry requires other steps described in the third paragraph of section 4 but not implemented here. The density change from first-step symmetry-preserved to self-consistent symmetry-broken is much smaller than the change from spherical to first-

step symmetry preserved. Importantly, the former density change yields an energy change for LSDA, PBE, and SCAN that is small compared to the error reductions in the sequence Hartree-Fock, LSDA, PBE, and SCAN. This finding quantifies a conclusion of Fertig and Kohn. These results have further strengthened our confidence that self-consistent symmetry breaking is the best way to calculate energies and energy differences. However, if we are interested in spatially symmetry-unbroken densities of atoms, then, as shown in this paper, they do not change the atomization energies significantly. Spin symmetries may also be preserved, if so desired, by starting from a spin-unpolarized equilibrium-ensemble, giving the Kohn-Sham potential all the symmetries of the external potential. Even fully symmetry-preserved densities are not claimed to be better than the self-consistent ones, except that they have the same symmetries as the exact densities (and their underlying Kohn-Sham non-interacting wavefunctions have the same symmetries as the true interacting wavefunctions). The "Differences" in our Tables I and V are not intended to be error estimates. What they suggest to us is that atomization energies that non-self-consistently preserve or self-consistently break the symmetries of the density are energetically close. Our conclusions about the non-empirical functionals LSDA, PBE, and SCAN are more firmly founded than our conclusions about symmetry preservation, where we have only taken a first step. However, that first step can also be the last step for B (one p-electron outside a closed subshell) and F (one p-hole in a closed shell), since for atoms in those electronic configurations only $L = 1$ is possible.

It was also found here that SCAN, which is usually accurate for atomization energies, underestimates that of strongly correlated singlet C_2 by about one electron volt. Spin symmetry breaking in the molecule might help, as it does for singlet Cr_2 [41], but was not found in our self-consistent calculation that started from a spin-unpolarized density. Starting the self-consistency cycle from a spin-polarized density is expected to improve the singlet C_2 atomization energy through spin symmetry breaking in SCAN, as it does for the PW91 GGA (which is very similar to PBE). Non-collinear spin-symmetry breaking and spatial symmetry breaking are also possible. It seems possible that the combination of a highly constrained functional like SCAN with full symmetry breaking might yield reliably accurate energetics.

Future work on the current topic might investigate symmetry breaking by SCAN in strongly-correlated singlet C_2, complete the symmetry preservation work for sp atoms by implementing the third paragraph of section 4, and extend the symmetry-preservation work to atoms with d and f electrons (where there are also issues of non-interacting pure-state v-representability of the exact ground-state density).

In the fifth chapter, we proposed that contrastive learning can be used as a flexible method to incorporate the uniform scaling property of DFT theory into the machine learn-

ing model design. And in principle, this method can be generalized to other exact physical conditions, such as rotational symmetry, spin scaling property, etc. Incorporating physical conditions into the machine learning model design can lead to significant reduced of the need of training data and provides insights into the machine learning XC functionals.

A similar effect occurs with human-designed density functionals: Those that are constructed to satisfy more exact constraints require fewer fit parameters that can be determined from smaller sets of molecular data, and a nonempirical meta-GGA functional [40] satisfying 17 exact constraints can perform rather well without any fitting to molecular data. The improvement of generalized gradient approximations (GGAs) or meta-GGAs by their global hybridization [265] with exact exchange is a good example, since the exact constraints on the underlying GGA or meta-GGA are preserved for any value of the fraction of exact exchange that is mixed with a complementary fraction of GGA or meta-GGA exchange.

APPENDIX A

A SINGLE ELECTRON IN A SPHERICALLY-SYMMETRIC POTENTIAL

Let us consider a single electron in a spherically-symmetric pseudopotential,

$$\tilde{V}(\mathbf{r}) = -Z/\mathbf{r} + \tilde{v}(\mathbf{r})$$

in its ground-state. Because the potentials are sphericals, the wavefunctions can be chosen to be spherical harmonics, i.e. s, p_x, p_y, p_z etc. Consequently, the general ground state wavefunction can be written as,

$$\psi(\mathbf{r}) = a_1\phi_{p_x}(\mathbf{r}) + a_2\phi_{p_y}(\mathbf{r}) + a_3\phi_{p_z}(\mathbf{r}) \tag{A.1}$$

Here, it will be proved by contradiction [1] that no choice of a_1, a_2, a_3 *(which can be complex or real)* can produce a spherical density. In other words, some reasonable-looking densities that are ensemble *v*-representable are not wavefunction *v*-representable.
Using Eq. (1.24) and keeping in mind that for us $N = 1$ and $q = 3$, we can expand Eq. (1.24) as,

$$\begin{aligned}(\delta_{11}d_1 - a_1^*a_1)n_{11}(\mathbf{r}) + (\delta_{12}d_2 - a_1^*a_2)n_{12}(\mathbf{r}) + (\delta_{13}d_3 - a_1^*a_3)n_{13}(\mathbf{r}) \\ + \delta_{21}d_1 - a_2^*a_1)n_{21}(\mathbf{r}) + (\delta_{22}d_2 - a_2^*a_2)n_{22}(\mathbf{r}) + (\delta_{23}d_3 - a_2^*a_3)n_{23}(\mathbf{r}) \\ + \delta_{31}d_1 - a_3^*a_1)n_{31}(\mathbf{r}) + (\delta_{32}d_2 - a_3^*a_2)n_{32}(\mathbf{r}) + (\delta_{33}d_3 - a_3^*a_3)n_{33}(\mathbf{r}) = 0\end{aligned} \tag{A.2}$$

where, $n(\mathbf{r}) = \sum_{i=1}^{3} d_i\, n_{ii}(\mathbf{r})$, and the probability of finding the electron in p_x, p_y or p_z results in $d_i = 1/3$ for a spherical ensemble density.

[1] Proof by Contradiction is a form of proof that establishes the truth or the validity of a proposition, by showing that assuming the proposition to be false leads to a contradiction.

As mentioned in Chapter 1, Ref. [15] assumed no linear dependencies (excluding those due to $n_{ij}(\mathbf{r}) \neq n_{ji}(\mathbf{r})$)). In this Appendix, we will show that for a special case, Levy's theorem is significantly true even when $n_{ij}(\mathbf{r}) = n_{ji}(\mathbf{r})$. Now, invoking $n_{ij}(\mathbf{r}) = n_{ji}(\mathbf{r})$ into Eq. (A.2) yields,

$$\begin{aligned}\Rightarrow (1/3 - a_1^* a_1) n_{11}(\mathbf{r}) - a_1^* a_2 n_{12}(\mathbf{r}) - a_1^* a_3 n_{13}(\mathbf{r}) \\ - a_2^* a_1 n_{12}(\mathbf{r}) + (1/3 - a_2^* a_2) n_{22}(\mathbf{r}) - a_2^* a_3 n_{23}(\mathbf{r}) \\ - a_3^* a_1 n_{13}(\mathbf{r}) - a_3^* a_2 n_{23}(\mathbf{r}) + (1/3 - a_3^* a_3) n_{33}(\mathbf{r}) = 0 \quad \text{(A.3)}\end{aligned}$$

$$\begin{aligned}\Rightarrow (1/3 - |a_1|^2) n_{11}(\mathbf{r}) - (a_1^* a_2 + a_2^* a_1) n_{12}(\mathbf{r}) - (a_1^* a_3 + a_3^* a_1) n_{13}(\mathbf{r}) + (1/3 - |a_2|^2) n_{22}(\mathbf{r}) \\ - (a_3^* a_2 + a_2^* a_3) n_{23}(\mathbf{r}) + (1/3 - |a_3|^2) n_{33}(\mathbf{r}) = 0 \quad \text{(A.4)}\end{aligned}$$

Let's write $a_i = r_i e^{i\theta_i}$
Then,

$$\begin{aligned}a_i^* a_j + a_j^* a_i &= r_i e^{-i\theta_i} r_j e^{i\theta_j} + r_j e^{-i\theta_j} r_i e^{i\theta_i} \\ &= 2 r_i r_j \cos(\theta_i - \theta_j)\end{aligned} \quad \text{(A.5)}$$

Using this to Eq. (A.4) yields,

$$\begin{aligned}(1/3 - |r_1|^2) n_{11}(\mathbf{r}) - 2 r_1 r_2 \cos(\theta_1 - \theta_2) n_{12}(\mathbf{r}) - 2 r_1 r_3 \cos(\theta_1 - \theta_3) n_{13}(\mathbf{r}) + (1/3 - |r_2|^2) n_{22}(\mathbf{r}) \\ - 2 r_2 r_3 \cos(\theta_2 - \theta_3) n_{23}(\mathbf{r}) + (1/3 - |r_3|^2) n_{33}(\mathbf{r}) = 0 \quad \text{(A.6)}\end{aligned}$$

The diagonal elements of n_{ii} give us three equations,

$$\begin{aligned}(1/3 - |r_1|^2) &= 0 \\ (1/3 - |r_2|^2) &= 0 \\ (1/3 - |r_3|^2) &= 0\end{aligned}$$

Solving these three equations we get,

$$\begin{aligned}|r_1| &= 1/\sqrt{3} \\ |r_2| &= 1/\sqrt{3} \\ |r_3| &= 1/\sqrt{3}\end{aligned}$$

Similarly, the other three equations can be read off as,

$$\begin{aligned}2 r_1 r_2 \cos(\theta_1 - \theta_2) &= 0 \\ 2 r_2 r_3 \cos(\theta_2 - \theta_3) &= 0 \\ 2 r_1 r_3 \cos(\theta_1 - \theta_3) &= 0\end{aligned}$$

Clearly, $2r_i r_j \neq 0$. So, all we have are,

$$\begin{aligned}\cos(\theta_1 - \theta_2) &= 0\\ \cos(\theta_2 - \theta_3) &= 0\\ \cos(\theta_1 - \theta_3) &= 0\end{aligned}$$

But, $\cos\alpha = 0$ when $\alpha = (2n+1)\pi/2\ \forall n$ where, $n \in \mathbb{Z}$ (set of integers),

Consequently,

$$\begin{aligned}\theta_1 - \theta_2 &= (2n_{12}+1)\frac{\pi}{2}\\ \theta_2 - \theta_3 &= (2n_{23}+1)\frac{\pi}{2}\\ \theta_1 - \theta_3 &= (2n_{13}+1)\frac{\pi}{2}\end{aligned}$$

where the three integers do not necessarily have to be equal.
Adding up the first two equations give us,

$$\theta_1 - \theta_3 = (2n_{12} + 2n_{23} + 2)\frac{\pi}{2}$$

Unfortunately, it directly contradicts the third equation. Here is how, let $\xi = n_{12} + n_{23}$ and equate the coefficients,

$$\begin{aligned}2\xi + 2 &= 2n_{13} + 1\\ \Rightarrow 2\xi + 1 &= 2n_{13}\\ \Rightarrow n_{13} &= \xi + 1/2\end{aligned}$$

However, by definition ξ is an integer. Hence, no choice of $\xi \in \mathbb{Z}$ makes n_{13} an integer according to the last expression. Rather, $n_{13} \in \mathbb{Q}$ (set of fractions). So, it directly contradicts our assumption. *Hence, there are no unique set of solutions for* $(\theta_1, \theta_2, \theta_3)$.

APPENDIX B

DERIVATION OF EQ. (2.7)

The total van der Waals energy can be divided into four components $E_{vdw}^{(1)}$, $E_{vdw}^{(2)}$, $E_{vdw}^{(3)}$, $E_{vdw}^{(4)}$: each comes from a different image potential V_{aan}, V_{abn}, V_{ban}, V_{bbn}. In turn, each of these potentials has four components, i.e., V_{aan} includes $V_{aan}^{(1)}$, $V_{aan}^{(2)}$, $V_{aan}^{(3)}$, and $V_{aan}^{(4)}$, where the first two components are the potentials at the positive charge and the last two components are for the negative charge.

Computing $\mathbf{V}_{aan}$

Now, $V_{aan}^{(1)}$ is the potential at the position of the charge at $(+q)$ due to $(+q)$,
Hence, $Z = -l/2$, $d = a + l/2$ and $q_0 = q$, using Equ. (2) of the main paper,

$$\begin{aligned} V_{aan}^{(1)} &= \frac{q_{aan}}{a + l/2 + (2n-2)\,c + a + l/2} \\ &= \frac{q_{aan}}{2a + l + (2n-2)\,c} \\ &= \frac{q_{aan}}{\Xi + l} \end{aligned} \tag{B.1}$$

where, $\Xi = 2a + (2n-2)c$
Using the Taylor's Expansion for $l << \Xi$,

$$V_{aan}^{(1)} \approx q_{aan}\left(1/\Xi - l/\Xi^2 + l^2/\Xi^3\right)$$

Now, $V_{aan}^{(2)}$ is the potential at the position of the charge at $(+q)$ due to $(-q)$,
Hence, $Z = -l/2$, $d = a - l/2$ and $q_0 = -q$, using Equ. (2) of the main paper,

$$
\begin{aligned}
V_{aan}^{(2)} &= \frac{q'_{aan}}{a+l/2+(2n-2)\,c+a-l/2} \\
&= \frac{q'_{aan}}{2a+(2n-2)\,c} \\
&= \frac{q'_{aan}}{\Xi}
\end{aligned}
\tag{B.2}
$$

V_{aan} at the position of the charge $(+q)$ is,

$$
\begin{aligned}
V_{aan}^{(\text{at } +q)} &= V_{aan}^{(1)} + V_{aan}^{(2)} \\
&= q_{aan}\left(1/\Xi - l/\Xi^2 + l^2/\Xi^3\right) + \left(\frac{q'_{aan}}{\Xi}\right)
\end{aligned}
\tag{B.3}
$$

But, $q_{aan} = \beta^{2n-1}(-q_0)$ where $q_0 = q$ and $q'_{aan} = \beta^{2n-1}(-q_0)$ where $q_0 = -q$
It is evident that $q'_{aan} = -q_{aan}$

Using these symmetry properties,

$$
V_{aan}^{(\text{at } +q)} = q_{aan}\left(-l/\Xi^2 + l^2/\Xi^3\right)
$$

Now, $V_{aan}^{(3)}$ is the potential at the position of the charge at $(-q)$ due to $(+q)$,
Hence, $Z = l/2$, $d = a + l/2$ and $q_0 = q$, using Equ. (2) of the main paper,

$$
\begin{aligned}
V_{aan}^{(3)} &= \frac{q_{aan}}{a-l/2+(2n-2)\,c+a+l/2} \\
&= \frac{q_{aan}}{2a+(2n-2)\,c} \\
&= \frac{q_{aan}}{\Xi}
\end{aligned}
\tag{B.4}
$$

Now, $V_{aan}^{(4)}$ is the potential at the position of the charge at $(-q)$ due to $(-q)$,
Hence, $Z = l/2$, $d = a - l/2$ and $q_0 = -q$, using Equ. (2) of the main paper,

$$
\begin{aligned}
V_{aan}^{(4)} &= \frac{q'_{aan}}{a-l/2+(2n-2)\,c+a-l/2} \\
&= \frac{q'_{aan}}{2a-l+(2n-2)\,c} \\
&= \frac{q'_{aan}}{\Xi - l}
\end{aligned}
\tag{B.5}
$$

Using the Taylor's Expansion for $l << \Xi$,

$$
V_{aan}^{(4)} \approx q'_{aan}\left(1/\Xi + l/\Xi^2 + l^2/\Xi^3\right)
$$

V_{aan} at the position of the charge $(-q)$ is,

$$\begin{aligned} V_{aan}^{(\text{at } -q)} &= V_{aan}^{(3)} + V_{aan}^{(4)} \\ &= q'_{aan}\left(1/\Xi + l/\Xi^2 + l^2/\Xi^3\right) + \left(\frac{q_{aan}}{\Xi}\right) \end{aligned} \tag{B.6}$$

Once again, using these symmetry properties, $q'_{aan} = -q_{aan}$

$$V_{aan}^{(\text{at } -q)} = q'_{aan}\left(l/\Xi^2 + l^2/\Xi^3\right)$$

Finally, total van der Waals energy due to the potential V_{aan} is,

$$\begin{aligned} E_{vdw}^{(1)} &= qV_{aan}^{(\text{at } +q)} - qV_{aan}^{(\text{at } -q)} \\ &= qq_{aan}\left(-l/\Xi^2 + l^2/\Xi^3\right) - qq'_{aan}\left(l/\Xi^2 + l^2/\Xi^3\right) \\ &= 2qq_{aan}\left(l^2/\Xi^3\right) \\ &= -2q^2\beta^{(2n-1)}\left(l^2/\Xi^3\right) \end{aligned} \tag{B.7}$$

Computing $\mathbf{V}_{abn}$

Now, $V_{abn}^{(1)}$ is the potential at the position of the charge at $(+q)$ due to $(+q)$,
Hence, $Z = -l/2$, $d = a + l/2$ and $q_0 = q$, using Equ. (3) of the main paper,

$$\begin{aligned} V_{abn}^{(1)} &= \frac{q_{abn}}{b + l/2 + (2n-1)\,c + a + l/2} \\ &= \frac{q_{abn}}{a + b + l + (2n-1)\,c} \\ &= \frac{q_{abn}}{\xi + l} \end{aligned} \tag{B.8}$$

where, $\xi = a + b + (2n-1)c$
Using the Taylor's Expansion for $l << \xi$,

$$V_{abn}^{(1)} \approx q_{abn}\left(1/\xi - l/\xi^2 + l^2/\xi^3\right)$$

Now, $V_{abn}^{(2)}$ is the potential at the position of the charge at $(+q)$ due to $(-q)$,
Hence, $Z = -l/2$, $d = a - l/2$ and $q_0 = -q$, using Equ. (3) of the main paper,

$$\begin{aligned} V_{abn}^{(2)} &= \frac{q'_{abn}}{b + l/2 + (2n-1)\,c + a - l/2} \\ &= \frac{q'_{abn}}{a + b + (2n-1)\,c} \\ &= \frac{q'_{abn}}{\xi} \end{aligned} \tag{B.9}$$

V_{abn} at the position of the charge $(+q)$ is,

$$\begin{aligned} V_{abn}^{(\text{at } +q)} &= V_{abn}^{(1)} + V_{abn}^{(2)} \\ &= q_{abn}\left(1/\xi - l/\xi^2 + l^2/\xi^3\right) + \left(\frac{q'_{abn}}{\xi}\right) \end{aligned} \tag{B.10}$$

It is evident that $q'_{abn} = -q_{abn}$ Using these symmetry properties,

$$V_{abn}^{(\text{at } +q)} = q_{abn}\left(-l/\xi^2 + l^2/\xi^3\right)$$

Now, $V_{abn}^{(3)}$ is the potential at the position of the charge at $(-q)$ due to $(+q)$, Hence, $Z = l/2$, $d = a + l/2$ and $q_0 = q$, using Equ. (3) of the main paper,

$$\begin{aligned} V_{abn}^{(3)} &= \frac{q_{abn}}{b - l/2 + (2n-1)\,c + a + l/2} \\ &= \frac{q_{abn}}{a + b + (2n-1)\,c} \\ &= \frac{q_{abn}}{\xi} \end{aligned} \tag{B.11}$$

Now, $V_{abn}^{(4)}$ is the potential at the position of the charge at $(-q)$ due to $(-q)$, Hence, $Z = l/2$, $d = a - l/2$ and $q_0 = -q$, using Equ. (3) of the main paper,

$$\begin{aligned} V_{abn}^{(4)} &= \frac{q'_{abn}}{b - l/2 + (2n-1)\,c + a - l/2} \\ &= \frac{q'_{abn}}{a + b - l + (2n-1)\,c} \\ &= \frac{q'_{abn}}{\xi - l} \end{aligned} \tag{B.12}$$

Using the Taylor's Expansion for $l << \xi$,

$$V_{abn}^{(4)} \approx q'_{abn}\left(1/\xi + l/\xi^2 + l^2/\xi^3\right)$$

V_{abn} at the position of the charge $(-q)$ is,

$$\begin{aligned} V_{abn}^{(\text{at } -q)} &= V_{abn}^{(3)} + V_{abn}^{(4)} \\ &= q'_{abn}\left(1/\xi + l/\xi^2 + l^2/\xi^3\right) + \left(\frac{q_{bbn}}{\xi}\right) \end{aligned} \tag{B.13}$$

Once again, using the symmetry, $q'_{abn} = -q_{abn}$

$$V_{abn}^{(\text{at } -q)} = q'_{abn}\left(l/\xi^2 + l^2/\xi^3\right)$$

Finally, total van der Waals energy due to the potential V_{abn} is,

$$\begin{aligned} E_{vdw}^{(2)} &= qV_{abn}^{(\text{at } +q)} - qV_{abn}^{(\text{at } -q)} \\ &= qq_{abn}\left(-l/\xi^2 + l^2/\xi^3\right) - qq'_{abn}\left(l/\xi^2 + l^2/\xi^3\right) \\ &= 2qq_{abn}\left(l^2/\xi^3\right) \\ &= -2q^2\beta^{2n}\left(l^2/\xi^3\right) \end{aligned} \tag{B.14}$$

Computing $\mathbf{V}_{ban}$

The derivation of $E_{vdw}^{(3)}$ is exactly the same as above. Except $q_{ban} = \beta^{2n} q_0$. Eventually,

$$E_{vdw}^{(3)} = 2q^2\beta^{2n}\left(l^2/\xi^3\right)$$

Computing $\mathbf{V}_{bbn}$

Now, $V_{bbn}^{(1)}$ is the potential at the position of the charge at $(+q)$ due to $(+q)$,
Hence, $Z = -l/2$, $d = a + l/2$ and $q_0 = q$, using Equ. (4) of the main paper,

$$\begin{aligned} V_{bbn}^{(1)} &= \frac{q_{bbn}}{b + l/2 + (2n-1)\,c + a + l/2} \\ &= \frac{q_{bbn}}{a + b + l + (2n-1)\,c} \\ &= \frac{q_{abn}}{\xi + l} \end{aligned} \tag{B.15}$$

where, $\xi = a + b + (2n-1)c$
Using the Taylor's Expansion for $l << \xi$,

$$V_{bbn}^{(1)} \approx q_{bbn}\left(1/\xi - l/\xi^2 + l^2/\xi^3\right)$$

Now, $V_{bbn}^{(2)}$ is the potential at the position of the charge at $(+q)$ due to $(-q)$,
Hence, $Z = -l/2$, $d = a - l/2$ and $q_0 = -q$, using Equ. (3) of the main paper,

$$\begin{aligned} V_{bbn}^{(2)} &= \frac{q'_{bbn}}{b + l/2 + (2n-1)\,c + a - l/2} \\ &= \frac{q'_{bbn}}{a + b + (2n-1)\,c} \\ &= \frac{q'_{bbn}}{\xi} \end{aligned} \tag{B.16}$$

V_{bbn} at the position of the charge $(+q)$ is,

$$\begin{aligned} V_{bbn}^{(\text{at } +q)} &= V_{bbn}^{(1)} + V_{bbn}^{(2)} \\ &= q_{bbn}\left(1/\xi - l/\xi^2 + l^2/\xi^3\right) + \left(\frac{q'_{bbn}}{\xi}\right) \end{aligned} \tag{B.17}$$

It is evident that $q'_{bbn} = -q_{bbn}$ Putting these symmetry properties,

$$V_{bbn}^{(\text{at } +q)} = q_{bbn}\left(-l/\xi^2 + l^2/\xi^3\right)$$

Now, $V_{bbn}^{(3)}$ is the potential at the position of the charge at $(-q)$ due to $(+q)$, Hence, $Z = l/2$, $d = a + l/2$ and $q_0 = q$, using Equ. (4) of the main paper,

$$\begin{aligned} V_{bbn}^{(3)} &= \frac{q_{bbn}}{b - l/2 + (2n-1)\,c + a + l/2} \\ &= \frac{q_{bbn}}{a + b + (2n-1)\,c} \\ &= \frac{q_{bbn}}{\xi} \end{aligned} \tag{B.18}$$

Now, $V_{bbn}^{(4)}$ is the potential at the position of the charge at $(-q)$ due to $(-q)$, Hence, $Z = l/2$, $d = a - l/2$ and $q_0 = -q$, using Equ. (3) of the main paper,

$$\begin{aligned} V_{bbn}^{(4)} &= \frac{q'_{bbn}}{b - l/2 + (2n-1)\,c + a - l/2} \\ &= \frac{q'_{bbn}}{a + b - l + (2n-1)\,c} \\ &= \frac{q'_{bbn}}{\xi - l} \end{aligned} \tag{B.19}$$

Using the Taylor's Expansion for $l << \Xi$,

$$V_{bbn}^{(4)} \approx q'_{bbn}\left(1/\xi + l/\xi^2 + l^2/\xi^3\right)$$

V_{bbn} at the position of the charge $(-q)$ is,

$$\begin{aligned} V_{bbn}^{(\text{at } -q)} &= V_{bbn}^{(3)} + V_{bbn}^{(4)} \\ &= q'_{bbn}\left(1/\xi + l/\xi^2 + l^2/\xi^3\right) + \left(\frac{q_{bbn}}{\xi}\right) \end{aligned} \tag{B.20}$$

Once again, using the symmetry, $q'_{bbn} = -q_{bbn}$

$$V'_{bbn}{}^{(\text{at } -q)} = q'_{bbn}\left(l/\xi^2 + l^2/\xi^3\right)$$

Finally, total van der Waals energy due to the potential V_{bbn} is,

$$\begin{aligned} E^{(4)}_{vdw} &= qV^{(\text{at } +q)}_{bbn} - qV^{(\text{at } -q)}_{bbn} \\ &= qq_{bbn}\left(-l/\xi^2 + l^2/\xi^3\right) - qq'_{bbn}\left(l/\xi^2 + l^2/\xi^3\right) \\ &= 2qq_{bbn}\left(l^2/\xi^3\right) \\ &= -2q^2\beta^{(2n-1)}\left(l^2/\xi^3\right) \end{aligned} \tag{B.21}$$

Total Energy

$$\begin{aligned} E^{total}_{vdw} = -2q^2\beta^{(2n-1)}\left(l^2/\Xi^3\right) - 2q^2\beta^{2n}\left(l^2/\xi^3\right) \\ + 2q^2\beta^{2n}\left(l^2/\xi^3\right) - 2q^2\beta^{(2n-1)}\left(l^2/\xi^3\right) \end{aligned}$$

where, $\Xi = 2a + (2n-2)c$ and $\xi = a + b + (2n-1)c$
And, finally plugging $b = a + c$ gives us $\xi = 2a + 2nc$

APPENDIX C

DERIVATION OF EQ. (3.11)

In this appendix, we present the derivation of Eq. ((3.11)). The starting point to derive Eq. (3.11) is the general form of the electric potential produced by the image charge residing inside the cylindrical medium. It is given by Eq. 3.6. The electrostatic interaction energy E^{static}_{dipole}, between the dipole and the cylinder, is given by Eq.3.10. With the expression for $q_m(k)$ in Eq. 3.4 (with $\varepsilon_v = 1$) and the condition $l \ll \rho$, Eq. 3.10 can be written as,

$$\begin{aligned} E^{static}_{dipole} &= \frac{-2q^2l^2}{\pi} \sum_{m=-\infty}^{\infty} \int_0^{\infty} dk \frac{\varepsilon - 1}{\varepsilon - h_m(k)} \frac{I_m(ka)}{K_m(ka)} [K_m^{'}(k\rho)]^2 \\ &= \frac{-2q^2l^2}{\pi} \sum_{m=-\infty}^{\infty} dk\, \xi_m(k,\rho)\, \frac{\varepsilon - 1}{\varepsilon - h_m(k)} \end{aligned} \tag{C.1}$$

where the function $\xi_m(k,\rho)$ is defined in Eq. 3.12.

When the radius of the cylinder becomes large, $a \to \infty$, the cylinder becomes a flat surface with infinite thickness. The problem becomes one in which an instantaneous dipole interacts with a flat surface of a solid. For this situation, when the distance D between the dipole and the surface is large, the vdW interaction energy is given by the Zaremba-Kohn formula as [209],

$$E^{ZK} = -\frac{1}{4\pi D^3} \int_0^{\infty} du\, \alpha_1(iu) \frac{\varepsilon(iu) - 1}{\varepsilon(iu) + 1} \tag{C.2}$$

where $\varepsilon(iu)$ is the dynamic dielectric constant of the surface, and $\alpha_1(iu)$ is the dynamic dipolar polarizability of the adsorbate. The time average of Eq. C.1 will lead to Eq. C.2 at large a and large D. With the asymptotic expressions for the modified Bessel functions [266], i.e., when $x \to \infty$, $I_m(x) ß e^x/\sqrt{2\pi x}$ and $K_m(x) ß e^{-x}\sqrt{/(2x)}$, it can be shown that $h_m(k) \to -1$ as $a \to \infty$. Besides, we have,

$$\lim_{a\to\infty} \frac{I_m(ka)}{K_m(ka)} \to \frac{e^{2ka}}{\pi} \tag{C.3}$$

$$\lim_{a\to\infty}[K_m^{'}(k\rho)]^2 \to \frac{\pi e^{-2k(D+a)}}{2k(D+a)} \tag{C.4}$$

When $a \to \infty$ and $D \to \infty$, Eq. C.1 becomes,

$$E_{dipole}^{static} \sim \frac{-2q^2l^2}{8\pi D^3}\frac{\varepsilon-1}{\varepsilon+1} \tag{C.5}$$

taking a form similar to that of Eq. C.2. The time average of Eq. C.5 can be written as,

$$\left\langle E_{dipole}^{static}\right\rangle_t \sim \frac{-2\left\langle q^2l^2\right\rangle_t}{8\pi D^3}\frac{\varepsilon-1}{\varepsilon+1} \tag{C.6}$$

where the symbol $\langle X\rangle_t$ represents the time average of quantity $X(t)$.

Since Eq. (A6) is equivalent to Eq. (C.2), this means that $\left\langle q^2l^2\right\rangle_t \frac{\varepsilon-1}{\varepsilon+1}$ is equivalent to

$$\int_0^\infty du\ \alpha_1(iu)\frac{\varepsilon(iu)-1}{\varepsilon(iu)+1}$$

. We generalize this equivalence to a finite a and D, and write the time average of Eq. (C.1) as Eq. (3.11) where $\alpha_1(iu)$ is defined after Eq. (3.13). Note that although ε in Eq. (C.1) is time-independent under the electrostatic limit, it should be understood as frequency-dependent, as shown in Eq. (3.11). The frequency dependent dielectric function is defined in the main text.

APPENDIX D

SUPPLEMENTARY TABLES OF CHAPTER 4

In this appendix, we present a comparison of our results with those of Ref. [12] using the same three functionals GGA-VWN [45, 46], GGA-PZ [45, 47], and LSDA-VWN [46]. Calculations reported in the Table are carried out with aug-cc-pvqz basis set.

Table D.1: Sym-P1 stands for non-spherical (first-step non-self-consistent spatial symmetry-preserving). Energies are in Hartree.

	GGA – VWN		**GGA – PZ**		**LSDA – VWN**	
Atom	**Our data**	**Ref. [12]**	**Our data**	**Ref. [12]**	**Our data**	**Ref. [12]**
B (Spherical)	-24.680	-24.681	-24.678	-24.679	-24.353	-24.353
B (Sym-P1)	-24.688	-24.687	-24.686	-24.685	-24.355	-24.354
C (Spherical)	-37.890	-37.891	-37.885	-37.887	-37.468	-37.469
C (Sym-P1)	-37.896	-37.896	-37.891	-37.891	-37.468	-37.468
O (Spherical)	-75.143	-75.146	-75.137	-75.140	-74.522	-74.523
O (Sym-P1)	-75.158	-75.159	-75.151	-75.152	-74.526	-74.526
F (Spherical)	-99.831	-99.838	-99.826	-99.832	-99.106	-99.111
F (Sym-P1)	-99.843	-99.847	-99.837	-99.841	-99.106	-99.109

Table D.2: The percentage change of SCAN atomization energies between the two finest grids available in NWChem. Sym-P1 stands for Non-Spherical (first-step non-self-consistent spatial symmetry-preserving). The percentage changes are calculated as

$$(x-y)/x \times 100\%$$

where, x is the SCAN energy using grid `'huge'` and y is the SCAN energy using grid `'xfine'`.

	Sym-P1	Spherical
B_2	0.25%	0.22%
C_2	-0.05%	-0.01%
O_2	0.08%	0.09%
F_2	0.17%	0.14%

APPENDIX E

MACHINE LEARNING GLOSSARY

In this appendix, I listed definitions of some common machine learning terms.

- **Accuracy:** Percentage of correct predictions made by the model.
- **Clustering:** Unsupervised grouping of data into buckets.
- **Convergence:** A state reached during the training of a model when the loss changes very little between each iteration.
- **Dimension:** Dimension for machine learning and data scientist is different from physics. Here, dimension of data means how many features you have in your data ocean (dataset).
- **Discriminative model:** A discriminative model is a model of the conditional probability $p(y|\mathbf{x})$ of the target y, given an observation $\mathbf{x}$.
- **Epoch:** An epoch describes the number of times the algorithm sees the entire data set.
- **Feature:** With respect to a dataset, a feature represents an attribute and value combination. The term feature has different definitions in different contexts.
- **Feature Vector:** A list of features describing an observation with multiple attributes.
- **Generative model:** A generative model is a statistical model of the joint probability distribution $p(\mathbf{x}, y)$ on given observable variable $\mathbf{x}$ and target variable y.
- **Hyperparameters:** Hyperparameters are higher-level properties of a model such as how fast it can learn (learning rate) or complexity of a model. The depth of trees in a Decision Tree or number of hidden layers in a Neural Networks are examples of hyper parameters.

- **Label:** The "answer" portion of an observation in supervised learning.
- **Learning Rate:** The size of the update steps to take during optimization loops like Gradient Descent.
- **Neural Networks:** Neural Networks are mathematical algorithms modeled after the brain's architecture, designed to recognize patterns and relationships in data.
- **Transfer Learning:** A machine learning method where a model developed for a task is reused as the starting point for a model on a second task.
- **Unsupervised Learning:** Training a model to find patterns in an unlabeled dataset (e.g. clustering).

APPENDIX F

LIST OF PUBLICATIONS

F.1 Published Articles in Peer-reviewed Journals

- Weiyi Gong, Tao Sun, Hexin Bai, **Shah Tanvir ur Rahman Chowdhury**, Peng Chu, Anoj Aryal, Jie Yu, Haibin Ling, John Perdew, Qimin Yan. Incorporation of density scaling constraint in density functional design via contrastive representation learning. `https://arxiv.org/abs/2205.15071 (2022).`

- **Shah Tanvir ur Rahman Chowdhury**, and John P. Perdew. Spherical vs. Non-spherical and Symmetry-preserving vs. Symmetry-breaking Densities of Open-shell Atoms in Density Functional Theory.
  Journal of Chemical Physics, 155, 234110 (2021).

- **Shah Tanvir ur Rahman Chowdhury**, Hong Tang, and John P. Perdew. van der Waals corrected Density Functionals for Cylindrical Surface: Ammonia and Nitrogen dioxide adsorbed on a Single-walled Carbon Nanotube.
  Physical Review B, 103(19), 195410 (2021).

- Hong Tang, **Shah Tanvir ur Rahman Chowdhury**, Jianmin Tao, and John P. Perdew. Density Functionals Combined with van der Waals Corrections for Graphene Adsorbed on Layered Materials.
  Physical Review B, 101(19), 195426 (2020).

- A. D. Kaplan, B. Santra, P. Bhattarai, K. Wagle, **Shah Tanvir ur Rahman Chowdhury**, P. Bhetwal, Jie Yu, Hong Tang, K. Burke, Mel Levy and John P. Perdew. Simple hydrogenic estimates for the exchange and correlation energies of atoms and atomic ions, with implications for density functional theory.
  Journal of Chemical Physics, 153, 074114 (2020).

F.2 Conference Presentation

Damped Zaremba-Kohn (dZK) model: van der Waals Corrected Density Functionals for Cylindrical Surfaces,
American Physical Society (APS) *March Meeting* March, 2021

REFERENCES

[1] Eberhard Engel and Reiner M Dreizler. *Density functional theory*. Springer, 2013.

[2] Max Born and Robert Oppenheimer. "Zur quantentheorie der molekeln". *Annalen der physik* 389.20 (1927), pp. 457–484.

[3] Douglas Rayne Hartree. "The wave mechanics of an atom with a non-coulomb central field. Part II. Some results and discussion". *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 24. 1. Cambridge University Press. 1928, pp. 111–132.

[4] Vladimir Fock. "Näherungsmethode zur Lösung des quantenmechanischen Mehrkö -rperproblems". *Zeitschrift für Physik* 61.1 (1930), pp. 126–148.

[5] Clemens Carel Johannes Roothaan. "New developments in molecular orbital theory". *Reviews of modern physics* 23.2 (1951), p. 69.

[6] Chr Møller and Milton S Plesset. "Note on an approximation treatment for many-electron systems". *Physical review* 46.7 (1934), p. 618.

[7] C David Sherrill and Henry F Schaefer III. "The configuration interaction method: Advances in highly correlated approaches". *Advances in quantum chemistry*. Vol. 34. Elsevier, 1999, pp. 143–269.

[8] Rodney J Bartlett. "Coupled-cluster approach to molecular structure and spectra: a step toward predictive quantum chemistry". *The Journal of Physical Chemistry* 93.5 (1989), pp. 1697–1708.

[9] Takao Komatsu. "On Tasoev's continued fractions". *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 134. 1. Cambridge University Press. 2003, pp. 1–12.

[10] Enrico Fermi. "Un metodo statistico per la determinazione di alcune priorieta dell'atome". *Rend. Accad. Naz. Lincei* 6.602-607 (1927), p. 32.

[11] Edward Teller. "On the stability of molecules in the Thomas-Fermi theory". *Reviews of Modern Physics* 34.4 (1962), p. 627.

[12] Paul AM Dirac. "Note on exchange phenomena in the Thomas atom". *Mathematical proceedings of the Cambridge philosophical society*. Vol. 26. 3. Cambridge University Press. 1930, pp. 376–385.

[13] CF v Weizsäcker. "Zur theorie der kernmassen". *Zeitschrift für Physik* 96.7-8 (1935), pp. 431–458.

[14] P Hohenberg and W. Kohn. "Density functional theory (DFT)". *Phys. Rev* 136 (1964), B864.

[15] Mel Levy. "Electron densities in search of Hamiltonians". *Physical Review A* 26.3 (1982), p. 1200.

[16] Laszlo Tisza. *Physics as Natural Philosophy: Essays in Honor of Laszlo Tisza on His Seventy-Fifth Birthday*. MIT Press, 1982.

[17] H Englisch and R Englisch. "Hohenberg-Kohn theorem and non-V-representable densities". *Physica A: Statistical Mechanics and its Applications* 121.1-2 (1983), pp. 253–268.

[18] Ernest Davidson. *Reduced density matrices in quantum chemistry*. Vol. 6. Elsevier, 2012.

[19] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". *Phys. Rev.* 140 (4A Nov. 1965), A1133–A1138.

[20] GL Oliver and JP Perdew. "Spin-density gradient expansion for the kinetic energy". *Physical Review A* 20.2 (1979), p. 397.

[21] Mel Levy and John P Perdew. "Hellmann-Feynman, virial, and scaling requisites for the exact universal density functionals. Shape of the correlation potential and diamagnetic susceptibility for atoms". *Physical Review A* 32.4 (1985), p. 2010.

[22] Mel Levy. "Density-functional exchange correlation through coordinate scaling in adiabatic connection and correlation hole". *Physical Review A* 43.9 (1991), p. 4637.

[23] GL Oliver and JP Perdew. "Spin-density gradient expansion for the kinetic energy". *Physical Review A* 20.2 (1979), p. 397.

[24] John P Perdew. "Size-consistency, self-interaction correction, and derivative discontinuity in density functional theory". *Advances in quantum chemistry*. Vol. 21. Elsevier, 1990, pp. 113–134.

[25] Elliott H Lieb and Stephen Oxford. "Improved lower bound on the indirect Coulomb energy". *International Journal of Quantum Chemistry* 19.3 (1981), pp. 427–439.

[26] Murray Gell-Mann and Keith A Brueckner. "Correlation energy of an electron gas at high density". *Physical Review* 106.2 (1957), p. 364.

[27] Rosemary A Coldwell-Horsfall and Alexei A Maradudin. “Zero-Point Energy of an Electron Lattice”. *Journal of Mathematical Physics* 1.5 (1960), pp. 395–404.

[28] David M Ceperley and Berni J Alder. “Ground state of the electron gas by a stochastic method”. *Physical review letters* 45.7 (1980), p. 566.

[29] John P Perdew and Yue Wang. “Accurate and simple analytic representation of the electron-gas correlation energy”. *Physical review B* 45.23 (1992), p. 13244.

[30] John P Perdew and Alex Zunger. “Self-interaction correction to density-functional approximations for many-electron systems”. *Physical Review B* 23.10 (1981), p. 5048.

[31] Puskar Bhattarai, Abhirup Patra, Chandra Shahi, and John P Perdew. “How accurate are the parametrized correlation energies of the uniform electron gas?” *Physical Review B* 97.19 (2018), p. 195128.

[32] David C Langreth and MJ Mehl. “Beyond the local-density approximation in calculations of ground-state electronic properties”. *Physical Review B* 28.4 (1983), p. 1809.

[33] Axel D Becke. “Density-functional exchange-energy approximation with correct asymptotic behavior”. *Physical review A* 38.6 (1988), p. 3098.

[34] Chengteh Lee, Weitao Yang, and Robert G Parr. “Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density”. *Physical review B* 37.2 (1988), p. 785.

[35] J.P. Perdew, K. Burke, and M. Ernzerhof. “Generalized Gradient Approximation Made Simple”. *Phys. Rev. Lett* 77 (1996), p. 3865.

[36] Rex W Godby, Michael Schlüter, and LJ Sham. “Self-energy operators and exchange-correlation potentials in semiconductors”. *Physical Review B* 37.17 (1988), p. 10159.

[37] John P Perdew and Adrienn Ruzsinszky. “Density-functional energy gaps of solids demystified”. *The European Physical Journal B* 91.6 (2018), pp. 1–5.

[38] John P Perdew, Adrienn Ruzsinszky, Gábor I Csonka, Lucian A Constantin, and Jianwei Sun. “Workhorse semilocal density functional for condensed matter physics and quantum chemistry”. *Physical Review Letters* 103.2 (2009), p. 026403.

[39] Lucian A Constantin, Eduardo Fabiano, and F Della Sala. “Meta-GGA exchange-correlation functional with a balanced treatment of nonlocality”. *Journal of chemical theory and computation* 9.5 (2013), pp. 2256–2263.

[40] J. Sun, A. Ruzsinszky, and J.P. Perdew. “Strongly Constrained and Appropriately Normed Semilocal Density”. *Functional, Phys. Rev. Lett* 115 (2015), p. 036402.

[41] Pál D Mezei, Gábor I Csonka, and Mihály Kállay. “Simple modifications of the SCAN meta-generalized gradient approximation functional”. *Journal of chemical theory and computation* 14.5 (2018), pp. 2469–2479.

[42] Thilo Aschebrock and Stephan Kümmel. “Ultranonlocality and accurate band gaps from a meta-generalized gradient approximation”. *Physical Review Research* 1.3 (2019), p. 033082.

[43] Albert P Bartók and Jonathan R Yates. “Regularized SCAN functional”. *J. Chem. Phys.* 150.16 (2019), p. 161101.

[44] D Mejıa-Rodrıguez and SB Trickey. “Comment on “Regularized SCAN functional”[J. Chem. Phys. 150, 161101 (2019)]”. *The Journal of chemical physics* 151.20 (2019), p. 207101.

[45] Albert P Bartók and Jonathan R Yates. “Response to “Comment on ‘Regularized SCAN functional’””. *J. Chem. Phys.* 151.20 (2019), p. 207102.

[46] Yoh Yamamoto, Alan Salcedo, Carlos M Diaz, Md Shamsul Alam, Tunna Baruah, and Rajendra R Zope. “Comparison of regularized SCAN functional with SCAN functional with and without self-interaction for a wide-array of properties”. *arXiv preprint arXiv:2004.13393* (2020).

[47] James W. Furness, Aaron D. Kaplan, Jinliang Ning, John P. Perdew, and Jianwei Sun. “Accurate and Numerically Efficient r2SCAN Meta-Generalized Gradient Approximation”. *The Journal of Physical Chemistry Letters* 11.19 (2020), pp. 8208–8215.

[48] Robert O Jones. “Density functional theory: Its origins, rise to prominence, and future”. *Reviews of modern physics* 87.3 (2015), p. 897.

[49] Jinbo Pan, Yan-Fang Zhang, Jingda Zhang, Huta Banjade, Jie Yu, Liping Yu, Shixuan Du, Adrienn Ruzsinszky, Zhenpeng Hu, and Qimin Yan. “Auxetic two-dimensional transition metal selenides and halides”. *npj Computational Materials* 6.1 (2020), pp. 1–6.

[50] Yan-Fang Zhang, Jinbo Pan, Huta Banjade, Jie Yu, Hsin Lin, Arun Bansil, Shixuan Du, and Qimin Yan. “Two-dimensional MX Dirac materials and quantum spin Hall insulators with tunable electronic and topological properties”. *Nano Research* 14.3 (2021), pp. 584–589.

[51] Carlo Adamo and Vincenzo Barone. “Toward reliable density functional methods without adjustable parameters: The PBE0 model”. *J. Chem. Phys.* 110.13 (1999), pp. 6158–6170.

[52] Jochen Heyd, Gustavo E Scuseria, and Matthias Ernzerhof. “Hybrid functionals based on a screened Coulomb potential”. *J. Chem. Phys.* 118.18 (2003), pp. 8207–8215.

[53] Jan Hermann, Robert A DiStasio Jr, and Alexandre Tkatchenko. “First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications”. *Chemical Reviews* 117.6 (2017), pp. 4714–4758.

[54] Max Dion, Henrik Rydberg, Elsebeth Schröder, David C Langreth, and Bengt I Lundqvist. “Van der Waals density functional for general geometries”. *Physical review letters* 92.24 (2004), p. 246401.

[55] Yingkai Zhang and Weitao Yang. “Comment on “Generalized gradient approximation made simple””. *Physical Review Letters* 80.4 (1998), p. 890.

[56] Kyuho Lee, Éamonn D Murray, Lingzhu Kong, I Lundqvist Bengt, and C Langreth David. “Higher-accuracy van der Waals density functional”. *Phys. Rev. B* 82.8 (2010), p. 081101.

[57] John P Perdew and Wang Yue. “Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation”. *Physical review B* 33.12 (1986), p. 8800.

[58] Éamonn D Murray, Kyuho Lee, and David C Langreth. “Investigation of exchange energy density functional accuracy for interacting molecules”. *Journal of Chemical Theory and Computation* 5.10 (2009), pp. 2754–2762.

[59] Jan Řezáč, Petr Jurečka, Kevin E Riley, Jiřı Černy, Haydee Valdes, Kristna Pluháčková, Karel Berka, Tomáš Řezáč, Michal Pitoňák, Jiřı Vondrášek, et al. “Quantum chemical benchmark energy and geometry database for molecular clusters and complex molecular systems (www. begdb. com): a users manual and examples”. *Collection of Czechoslovak Chemical Communications* 73.10 (2008), pp. 1261–1270.

[60] Handan Yildirim, Thomas Greber, and Abdelkader Kara. “Trends in adsorption characteristics of benzene on transition metal surfaces: Role of surface chemistry and van der Waals interactions”. *The Journal of Physical Chemistry C* 117.40 (2013), pp. 20572–20583.

[61] Klime, David R Bowler, and Angelos Michaelides. “Chemical accuracy for the van der Waals density functional”. *Journal of Physics: Condensed Matter* 22.2 (2009), p. 022201.

[62] O.A. Vydrov and T.Van Voorhis. “Nonlocal van der Waals density functional: The simpler the better”. *J. Chem. Phys* 133 (2010), p. 244103.

[63] R. Sabatini, T. Gorni, and S. de Gironcoli. “Nonlocal van der Waals density functional made simple and efficient”. *Phys. Rev. B* 87 (2013), p. 041108.

[64] Guillermo Román-Pérez and José M Soler. “Efficient implementation of a van der Waals density functional: application to double-wall carbon nanotubes”. *Physical review letters* 103.9 (2009), p. 096102.

[65] Herbert B. Callen and Theodore A. Welton. “Irreversibility and Generalized Noise”. *Phys. Rev.* 83 (1 1951), pp. 34–40.

[66] Adrienn Ruzsinszky, John P Perdew, and Gábor I Csonka. “The RPA atomization energy puzzle”. *Journal of chemical theory and computation* 6.1 (2010), pp. 127–134.

[67] Xinguo Ren, Patrick Rinke, Christian Joas, and Matthias Scheffler. “Random-phase approximation and its applications in computational chemistry and materials science”. *Journal of Materials Science* 47.21 (2012), pp. 7447–7471.

[68] Adrienn Ruzsinszky, John P Perdew, and Gábor I Csonka. “A simple but fully nonlocal correction to the random phase approximation”. *The Journal of chemical physics* 134.11 (2011), p. 114110.

[69] Carla P. Gomes, Bart Selman, and John M. Gregoire. “Artificial intelligence for materials discovery”. *MRS Bulletin* 44.7 (2019), pp. 538–544.

[70] Phuc H Le-Khac, Graham Healy, and Alan F Smeaton. “Contrastive representation learning: A framework and review”. *IEEE Access* 8 (2020), pp. 193907–193934.

[71] Henry W Lin, Max Tegmark, and David Rolnick. “Why does deep and cheap learning work so well?” *Journal of Statistical Physics* 168.6 (2017), pp. 1223–1247.

[72] David J. Tozer, Victoria E. Ingamells, and Nicholas C. Handy. “Exchange-correlation potentials”. *The Journal of Chemical Physics* 105.20 (1996), pp. 9200–9213.

[73] Christopher W. Murray, Nicholas C. Handy, and Gregory J. Laming. “Quadrature schemes for integrals of density functional theory”. *Molecular Physics* 78.4 (1993), pp. 997–1014.

[74] John C. Snyder, Matthias Rupp, Katja Hansen, Klaus-Robert Müller, and Kieron Burke. “Finding Density Functionals with Machine Learning”. *Phys. Rev. Lett.* 108 (25 June 2012), p. 253002.

[75] James Kirkpatrick, Brendan McMorrow, David H. P. Turban, Alexander L. Gaunt, James S. Spencer, Alexander G. D. G. Matthews, Annette Obika, Louis Thiry, Meire Fortunato, David Pfau, Lara Román Castellanos, Stig Petersen, Alexander W. R. Nelson, Pushmeet Kohli, Paula Mori-Sánchez, Demis Hassabis, and Aron J. Cohen. "Pushing the frontiers of density functionals by solving the fractional electron problem". *Science* 374.6573 (2021), pp. 1385–1389.

[76] John P Perdew and Karla Schmidt. "Jacob's ladder of density functional approximations for the exchange-correlation energy". *AIP Conference Proceedings*. Vol. 577. 1. American Institute of Physics. 2001, pp. 1–20.

[77] Sebastian Dick and Marivi Fernandez-Serra. "Using differentiable programming to obtain an energy and density-optimized exchange-correlation functional". *arXiv preprint arXiv:2106.04481* (2021).

[78] J.D. van der Waals and Ph D. "Dissertation, Over de continuiteit van den gasen vloeistoftoestand (On the Continuity of the Gaseous and Liquid States)". Universiteit Leiden, Netherlands, 1873.

[79] V.A. Parsegian. *Van der Waals Forces: A Handbook for Biologists, Chemists, Engineers, and Physicists*. Cambridge University Press, Cambridge, 2005.

[80] *Reviews in Computational Chemistry*. Ed. by K.B. Lipkowitz and T. R. Vol. 26. New York: Cundari (Wiley, 2008.

[81] M.J. Gillan, D. Alfè, A. Michaelides, and J. *Chem. Phys* 144 (2016), p. 130901.

[82] *Proceedings of the Symposium on Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials*. Ed. by K.M. Kadish and R.S. Ruoff. The Electrochemical Society, Pennington, New Jersey, 1997.

[83] J. Tao, John P. Perdew, and A. Ruzsinszky. *Phys. Rev. B* 81 (2010), p. 233102.

[84] D. Leckband, J. Israelachvili, and Q. *Rev. Biophys* 34.105 (2001).

[85] J. Zou, Z. Marcet, A.W. Rodriguez, M.T.H. Reid, A.P. McCauley, I.I. Kravchenko, T. Lu, Y. Bao, S.G. Johnson, and H. B. *Chan, Nat. Commun* 4.1845 (2013).

[86] K.A. Milton and Nat. *Photonics* 11.73 (2017).

[87] A.E. Nel, L. Mädler, D. Velegol, T. Xia, E.M. Hoek, P. Somasundaran, F. Klaessig, V. Castranova, and M. Thompson. *Nat. Mater* 8.543 (2009).

[88] B.B. Machta, S.L. Veatch, and J. Sethna. *Phys. Rev. Lett* 109 (2012), p. 138101.

[89] K. Bradonjic, J.D. Swain, A. Widom, Y.N. Srivastava, and J. *Phys.: Conf. Ser* 161 (2009), p. 012035.

[90] K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, and A. A. *Firsov, Science* 306.666 (2004).

[91] C. Cheng, Z. Li, A. Hambarde, P.B. Deotare, and A.C.S. Appl. *Mater. Interfaces* 10 (2018), p. 39336.

[92] S. Das, D. Pandey, J. Thomas, T. Roy, and Adv. *Mater* 31.1802722 (2019).

[93] P.K. Yang and C.P. Lee. *2D-layered nanomaterials for energy harvesting and sensing applications, in Applied Electromechanical Device and Machines for Electric Mobility Solutions (IntechOpen*. London, 2019.

[94] N. Choudhary, M.A. Islam, J.H. Kim, T. Ko, A. Schropp, L. Hurtado, D. Weitzman, L. Zhai, and Y. Jung. *Nanotoday* 19.16 (2018).

[95] J. l. Ning, J.W. Furness, Y. Zhang, A.C. Thenuwara, R.C. Remsing, M.L. Klein, D.R. Strongin, and J. Sun. *J. Catalysis* 374 (2019), p. 143.

[96] M.S. Stark, K.L. Kuntz, S.J. Martens, and S.C. Warren. *Adv. Mater* 31.1808213 (2019).

[97] Z. Tu and M. Wu. *Adv. Electr. Mater* 5.1800960 (2019).

[98] W. Han, R.K. Kawakami, M. Gmitra, and J. Fabian. *Nat. Nanotechnology 9* 794 (2014).

[99] J. Sun, Y. Choi, Y.J. Choi, S. Kim, J. Park, S. Lee, and J.H. Cho. *Adv. Mater* 14.1803831 (2019).

[100] J. Wang, F. Ma, M. Sun, and R.S.C. Adv. 2017.

[101] A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai. *Nature (London)* 424.654 (2003).

[102] C.M. Yang, Y.J. Kim, M. Endo, H. Kanoh, M. Yudasaka, S. Iijima, K. Kaneko, and J. *Am. Chem. Soc* 129.20 (2007).

[103] L. Kou, Y. Ma, C. Tang, Z. Sun, A. Du, and C. Chen. 2016.

[104] L. Yu, A. Ruzsinszky, and John P. Perdew. 2016.

[105] F. Bussolotti, H. Kawai, Z.E. Ooi, V. Chellappan, D. Thian, A.L.C. Pang, and K.E. J Goh. *Nano Futures* 2 (2018), p. 032001.

[106] J.A. Reyes-Retana and F. Cervantes-Sodi. *Sci. Rep.* 6.24093 (2016).

[107] C. Janisch, H. Song, C. Zhou, Z. Lin, A.L. Elías, D. Ji, M. Terrones, Q. Gan, and Z. Liu. *2D Materials*. Vol. 3. 2016, p. 025017.

[108] S. Srinivasan, G. Balasubramanian, and Langmuir. 2018.

[109] P.A. Alekseev, B.R. Borodin, F.A. Benimetskiy, A.N. Smirnov, V.Y. Davydov, S.P. Lebedev, A.A. Lebedev, M.S. Dunaevskiy, and J. *Phys. Conf. Series* 1092 (2018), p. 012002.

[110] O. Gunnarsson and B. I Lundqvist. *Phys. Rev. B* 13 (1976), p. 4274.

[111] H. Tang, M. Chiu, C. Tseng, S. Yang, K. Hou, S. We, J. Huang, Y. Lin, C. Lien, and L. Li. *ACS Nano* 11 (2017), p. 12817.

[112] F. Ye, J. Lee, and P.X. Feng. *Nanoscale 9* 18208 (2017).

[113] A. Azizi, S. Eichfeld, G. Geschwind, K. Zhang, B. Jiang, D. Mukherjee, L. Hossain, A.F. Piasecki, B. Kabius, J.A. Robinson, and N. Alem. *ACS Nano 9* 4882 (2015).

[114] D.C. Langreth and John P. Perdew. *Phys. Rev. B* 15 (1977), p. 2884.

[115] D. Pines and D. Bohm. *Phys. Rev. D.* 85.338 (1952).

[116] D. Bohm and D. Pines. *Phys. Rev* 92.609 (1953).

[117] Lundqvist. *Phys. Rev. Lett* 92 (2004), p. 246401.

[118] K. Lee, E.D. Murray, L. Kong, B.I. Lundqvist, and D. C Langreth. *Phys. Rev. B* 82 (2010), p. 081101.

[119] V. R. Cooper. *Phys. Rev. B* 81 (2010), p. 161104.

[120] W. Kohn and L.J. Sham. *Phys. Rev.* 140 (1965), p. 1133.

[121] J. Klimeš, D.R. Bowler, and A. Michaelides. *J. Phys. Condens. Matter* 22 (2010), p. 022201.

[122] O.A. Vydrov and T.Van Voorhis. *J. Chem. Phys* 133 (2010), p. 244103.

[123] R. Sabatini, T. Gorni, and S. de Gironcoli. *Phys. Rev. B* 87 (2013), p. 041108.

[124] S. Grimme. *J. Comput. Chem* 27.1787 (2006).

[125] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. *J. Chem. Phys* 132 (2010), p. 154104.

[126] A. Tkatchenko and M. Scheffler. *Phys. Rev. Lett* 102.73005 (2009).

[127] A. Tkatchenko, R.A. DiStasio Jr., R. Car, and M. Scheffler. *Phys. Rev. Lett* 108.236402 (2012).

[128] W. Liu, F. Maaß, M. Willenbocke, C. Bronner, M. Schulze, S. Soubatch, F.S. Tautz, P. Tegeder, and A. Tkatchenko. *Phys. Rev. Lett* 115 (2015), p. 036104.

[129] V.G. Ruiz, W. Liu, E. Zojer, M. Scheffler, and A. Tkatchenko. *Phys. Rev. Lett* 108 (2012), p. 146103.

[130] E. Zaremba and W. Kohn. *Phys. Rev. B* 13 (1976), p. 2270.

[131] J.N. Israelachvili. *Intermolecular and Surface Forces (Academic*. Cambridge, MA, 2011.

[132] J.F. Dobson, A. White, and A. Rubio. *Phys. Rev. Lett* 96 (2006), p. 073201.

[133] X.-P. Jiang, F. Toigo, and M. W Cole. *Surf. Sci* 145.281 (1984).

[134] X.-P. Jiang, F. Toigo, and M. W Cole. *Surf. Sci* 148.21 (1984).

[135] J. Tao, H. Tang, A. Patra, P. Bhattarai, and John P. Perdew. *Phys. Rev. B 97* 165403 (2018).

[136] H. Tang, J. Tao, A. Ruzsinszky, and John P. Perdew. *J. Phys. Chem. C* 123 (2019), p. 13748.

[137] S. Adhikari, H. Tang, B. Neupane, A. Ruzsinszky, and G. I Csonka. *Phys. Rev. Mater* 4 (2020), p. 025005.

[138] T. Sometani. *J. Eur. Phys* 21.549 (2000).

[139] A.D. McLachlan. *Mol. Phys* 7.381 (1964).

[140] F.O. Kannemann and A.D. Becke. *J. Chem. Phys* 136 (2012), p. 034109.

[141] Jr. R. A. Breckenridge R. W. Shaw and A Sher. *Phys. Rev. B* 10 (1974), p. 2483.

[142] Jr. R. A. Breckenridge R. W. Shaw and A. Sher. "Improved calculations of the complex dielectric constant of semiconductors". *Phys. Rev. B* 10 (1974), p. 2483.

[143] F. Zheng, J. Tao, and A.M. Rappe. "Frequency-dependent dielectric function of semiconductors with application to physisorption". *Phys. Rev. B 95* 35203 (2017).

[144] D.R. Penn. "Wave-number-dependent dielectric function of semiconductors". *Phys. Rev* 128.2093 (1962).

[145] W. Lu, D. Wang, and L. Chen. "Near-static dielectric polarization of individual carbon nanotubes, Nano Lett". *7* 2729 (2007).

[146] F. Zheng, J. Tao, and A. M Rappe. *Phys. Rev. B. 95* 35203 (2017).

[147] A. Kumar and P. K Ahluwalia. *Physica B* 407 (2012), p. 4627.

[148] *See Supplemental Material*. URL: `http://link.aps.org/`.

[149] T. Björkman, A. Gulans, A.V. Krasheninnikov, and R. M Nieminen. *Phys. Rev. Lett* 108.235502 (2012).

[150] T. Björkman. *J.Chem. Phys* 141 (2014), p. 074708.

[151] G. Kresse and J. Furthmüller. *Phys. Rev.B* 54 (1996), p. 11169.

[152] P. E. Blöchl. *Phys. Rev. B* 50 (1994), p. 17953.

[153] G. Kresse and D. Joubert. *Phys. Rev. B* 59.1758 (1999).

[154] H. Peng and J. Perdew. *Phys. Rev. B 95* 81105 (2017).

[155] Y. Ma, Y. Dai, M. Guo, C. Niu, and B. Huang. *Nanoscale* 3.3883 (2011).

[156] Z. Wang, Q. Chen, and J. Wang. *J. Phys. Chem. C* 119 (2015), p. 4752.

[157] J. Jiang and H.S. Park. *Appl. Phys. Lett* 105 (2014), p. 033108.

[158] T.P. Kaloni, L. Kou, T. Frauenheim, and U. Schwingenschlögl. *Appl. Phys. Lett* 105.233112 (2014).

[159] S. Li and C. Zhang. *J. Appl. Phys* 114 (2013), p. 183709.

[160] H. Eshuis, J.E. Bates, and F. Furche. *Theor. Chem. Acc* 131.1084 (2012).

[161] T. Olsen and K. S Thygesen. *Phys. Rev. B* 87 (2013), p. 075111.

[162] T. Gould, S. Lebègue, J.G. Ángyán, and T. Bucko. *J.Chem. Theory Comput* 12.5920 (2016).

[163] M. Kim, W.J. Kim, T. Gould, E.K. Lee, S. Lebègue, and H. Kim. *J.Am. Chem. Soc* 142.2346 (2020).

[164] T. Olsen, C.E. Patrick, J.E. Bates, A. Ruzsinszky, and K. S Thygesen. *npj Comput. Mater* 5.106 (2019).

[165] T. Olsen and K. S Thygesen. *Phys. Rev. B* 88115131 (2013).

[166] B. Ramberger, T. Schäfer, and G. Kresse. *Phys. Rev. Lett* 118.106403 (2017).

[167] H. Tang, S.T. u. R. Chowdhury, J. Tao, and J.P. Perdew. "Density functionals combined with van der Waals corrections for graphene adsorbed on layered materials". *Phys. Rev. B* 101 (1954).

[168] Alberto Ambrosetti, Anthony M. Reilly, Robert A. DiStasio, and Alexandre Tkatchenko. "Long-range correlation energy calculated from coupled atomic response functions". *The Journal of Chemical Physics* 140.18 (2014), 18A508.

[169] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A.Dal Corso, S. Fabris, G. Fratesi, S. de Gironcoli, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, and R.M. Wentzcovitch. *J.Phys.: Condens. Matter* 21.395502 (2009).

[170] A. Ambrosetti, P.L. Silvestrelli, and A. Tkatchenko. *Phys. Rev. B 95* 235417 (2017).

[171] B. Sachs, T.O. Wehling, M.I. Katsnelson, and A. I Lichtenstein. *Phys. Rev. B* 84 (1954).

[172] A. Ambrosetti and P. L Silvestrelli. *Carbon* 139.486 (2018).

[173] V.A. Parsegian. *Van der Waals Forces: A Handbook for Biologists, Chemists, Engineers, and Physicists*. Cambridge University Press, Cambridge, 2005.

[174] *Reviews in Computational Chemistry, Volume 26*. Ed. by K.B. Lipkowitz, T.R. Cundari, and D.B. Boyd. New York: John Wiley Sons, 2008.

[175] M.J. Gillan, D. Alfè, and A. Michaelides. "Perspective: How good is DFT for water?" *J. Chem. Phys* 144 (2016), p. 130901.

[176] Jr. J. Hermann R. A. DiStasio and A. Tkatchenko. "Firstprinciples models for van der Waals interactions in molecules and materials: Concepts, theory, and applications". *Chem. Rev* 117 (2017), p. 4714.

[177] J. Tao, J.P. Perdew, A. Ruzsinszky, and Long-range van der. "Waals attraction and alkali-metal lattice constants". *Phys. Rev. B* 81 (2010), p. 233102.

[178] O.A.von Lilienfeld R. A. DiStasio Jr. and A. Tkatchenko. "Collective many-body van der Waals interactions in molecular systems". *Proc. Natl. Acad. Sci. USA* 109 (2012), p. 14791.

[179] K. Autumn, Y.A. Liang, S.T. Hsieh, W. Zesch, W.P. Chan, T.W. Kenny, R. Fearing, and R.J. Full. "Adhesive force of a single gecko foot-hair". *Nature* 405.681 (2000).

[180] K. Autumn, M. Sitti, Y.A. Liang, A.M. Peattie, W.R. Hansen, S. Sponberg, T.W. Kenny, R. Fearing, J.N. Israelachvili, and R.J. Full. "Evidence for van der Waals adhesion in gecko setae". *Proc. Natl. Acad. Sci. USA* 99.12252 (2002).

[181] B. Rozitis, E. MacLennan, and J.P. Emery. "Cohesive forces prevent the rotational breakup of rubble-pile asteroid (29075) 1950". *DA, Nature* 512.174 (2014).

[182] D.J. Scheeres, C.M. Hartzell, P. Sánchez, and M. Swift. "Scaling forces to asteroid surfaces: The role of cohesion". *Icarus* 210 (2010), p. 968.

[183] O. Gunnarsson and B.I. Lundqvist. "Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism". *Phys. Rev. B* 13 (1976), p. 4274.

[184] D.C. Langreth and J.P. Perdew. "Exchange-correlation energy of a metallic surface: Wave-vector analysis". *Phys. Rev. B* 15 (1977), p. 2884.

[185] D. Pines and D. Bohm. "A collective description of electron interactions: II. Collective vs individual particle aspects of the interactions". *Phys. Rev* 85.338 (1952).

[186] D. Bohm and D. Pines. "A collective description of electron interactions: IV. Electron interaction in metals". *Phys. Rev* 92.609 (1953).

[187] "Van der Waals Density Functional for". *General Geometries, Phys. Rev. Lett* 92 (2004), p. 246401.

[188] K. Lee, E.D. Murray, L. Kong, B.I. Lundqvist, and D.C. Langreth. “Higher-accuracy van der Waals density”. *functional, Phys. Rev. B* 82 (2010), p. 081101.

[189] V.R. Cooper and Van der. “Waals density functional: An appropriate exchange functional”. *Phys. Rev. B* 81 (2010), p. 161104.

[190] J. Klimeš, D.R. Bowler, and A. Michaelides. “Chemical accuracy for the van der Waals density”. *functional, J. Phys.: Condens. Matter* 22 (2010), p. 022201.

[191] J. Klimeš, D.R. Bowler, and A. Michaelides. “Van der Waals density functionals applied to solids”. *Phys. Rev. B* 83 (1951).

[192] I. Hamada and van der. “Waals density functional made accurate”. *Phys. Rev. B* 89 (2014), p. 121103.

[193] K. Berland and P. Hyldgaard. “Exchange functional that tests the robustness of the plasmon description of the van der Waals density”. *functional, Phys. Rev. B* 89 (2014), p. 035412.

[194] S. Grimme. “Semiempirical GGA-type density functional constructed with a long-range dispersion correction”. *J. Comput. Chem* 27.1787 (2006).

[195] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. “A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu”. *J. Chem. Phys* 132 (2010), p. 154104.

[196] A. Tkatchenko and M. Scheffler. “Accurate Molecular van der Waals Interactions from Ground-State Electron Density and Free-Atom Reference”. *Data, Phys. Rev. Lett* 102 (2009), p. 073005.

[197] A. Tkatchenko, R.Car R. A. DiStasio Jr., and M. Scheffler. “Accurate and Efficient Method for Many-Body van der Waals Interactions”. *Phys. Rev. Lett* 108.236402 (2012).

[198] W. Liu, F. Ma, M. Willenbocke, C. Bronner, M. Schulze, S. Soubatch, F.S. Tautz, P. Tegeder, and A. Tkatchenko. “Quantitative Prediction of Molecular Adsorption: Structure and Binding of Benzene on Coinage Metals”. *Phys. Rev. Lett* 115 (2015), p. 036104.

[199] S. Iijima. “Helical microtubules of graphitic carbon”. *Nature* 354.56 (1991).

[200] M.S. Dresselhaus, G. Dresselhaus, and P.C. Eklund. *Science of Fullerenes and Carbon Nanotube: Their Properties and Applications*. Elsevier, Amsterdam, 1996.

[201] S.J. Tans, M.H. Devoret, H. Dai, A. Thess, R.E. Smalley, L. Geerligs, and C. Dekker. “Individual single-wall carbon nanotubes as quantum wires”. *Nature* 386 (1997), p. 474.

[202] R.H. Baughman, A.A. Zakhidov, and W.A. de Heer. “Carbon nanotubes—the route toward applications”. *Science* 297.787 (2002).

[203] S.J. Tans, A.R. Verschueren, and C. Dekker. “Roomtemperature transistor based on a single carbon nanotube”. *Nature* 393.49 (1998).

[204] J. Kong, N.R. Franklin, C. Chou, M.G. Chaplin, S. Peng, K. Cho, and H. Dai. “Nanotube molecular wires as chemical sensors”. *Science* 287.622 (2000).

[205] H. Chang, J.D. Lee, S.M. Lee, and Y.H. Lee. “Adsorption of NH3 and NO2 molecules on carbon nanotubes”. *Appl. Phys. Lett. 79* 3863 (2001).

[206] W. Kohn and L.J. Sham. “Self-consistent equations including exchange and correlation effects”. *Phys. Rev* 140 (1965), p. 1133.

[207] T. Emig, R.L. Jaffe, M. Kardar, and A. Scardicchio. “Casimir Interaction between a Plate and a Cylinder”. *Phys. Rev. Lett* 96 (2006), p. 080403.

[208] R.F. Rajter, R. Podgornik, V.A. Parsegian, R.H. French, and W. Y. “Ching, van der Waals–London dispersion interactions for optically anisotropic cylinders: Metallic and semiconducting single-wall carbon nanotubes”. *Phys. Rev. B* 76 (2007), p. 045417.

[209] E. Zaremba and W. Kohn. “Van der Waals interaction between an atom and a solid surface”. *Phys. Rev. B* 13 (1976), p. 2270.

[210] J. Tao, H. Tang, A. Patra, P. Bhattarai, and J.P. Perdew. “Modeling the physisorption of graphene on metals”. *Phys. Rev. B 97* 165403 (2018).

[211] H. Tang, J. Tao, A. Ruzsinszky, J.P. Perdew, and van der. “Waals correction to the physisorption of graphene on metal surfaces”. *J. Phys. Chem. C* 123 (2019), p. 13748.

[212] S. Adhikari, H. Tang, B. Neupane, A. Ruzsinszky, and G.I. Csonka. “Molecule-surface interaction from van der Waals-corrected semilocal density functionals: The example of thiophene on transition-metal surfaces”. *Phys. Rev. Materials* 4 (2020), p. 025005.

[213] J.A. Hernandes and A.K.T. Assis. “Electric potential due to an infinite conducting cylinder with internal or external point charge”. *J. Electrostatics* 63.1115 (2005).

[214] S.T. Cui. “Electrostatic potential in cylindrical dielectric media using the image charge method, Mol”. *Phys* 104.2993 (2006).

[215] L.X. Benedict, S.G. Louie, and M.L. Cohen. “Static polarizabilities of single-wall carbon nanotubes”. *Phys. Rev. B* 52 (1995), p. 8541.

[216] B. Kozinsky and N. Marzari. “Static Dielectric Properties of Carbon Nanotubes from First Principles”. *Phys. Rev. Lett* 96.166801 (2006).

[217] J. Tao and A.M. Rappe. “Physical Adsorption: Theory of van der Waals Interactions between Particles and Clean Surfaces”. *Phys. Rev. Lett* 112 (2014), p. 106101.

[218] G. Vidali and M.W. Cole. “Long range interaction between a He atom and a semiconductor surface, Surf”. *Sci* 107 (1981), p. 374.

[219] R.D. Johnson, ed. *Comparison and Benchmark Database, NIST Standard Reference Database No. 101, Release 20*. III. URL: `http://cccbdb.nist.gov/`.

[220] G. Kresse and J. Furthmüller. “Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set”. *Phys. Rev. B* 54 (1996), p. 11169.

[221] P.E. Blöchl. “Projector augmented-wave method”. *Phys. Rev. B* 50 (1994), p. 17953.

[222] G. Kresse and D. Joubert. “From ultrasoft pseudopotentials to the projector augmented-wave method”. *Phys. Rev. B* 59.1758 (1999).

[223] Y. Lia, M. Hodak, W. Lu, and J. Bernholc. “Mechanisms of NH3 and NO2 detection in carbon-nanotube-based sensors: An ab initio investigation”. *Carbon* 101.177 (2016).

[224] S. Peng, K. Cho, P. Qi, and H. Dai. “Ab initio study of CNT NO2 gas sensor”. *Chem. Phys. Lett* 387 (2004), p. 271.

[225] L. Bai and Z. Zhou. “Computational study of B-or N-doped single-walled carbon nanotubes as NH3 and NO2 sensors”. *Carbon* 45 (2007), p. 2105.

[226] J. Wiktor, F. Ambrosio, and A. Pasquarello. “Assessment of the SCAN+ rVV10 functional for the structure of liquid water”. *J. Chem. Phys* 147 (2017), p. 216101.

[227] J. Yu, G. Fiorin, H. Peng, M.L. Klein, and J.P. Perdew. “Different bonding type along each crystallographic axis: Computational study of poly (p-phenylene terephthalamide)”. *Phys. Rev. Materials* 4 (2020), p. 055601.

[228] M. Kim, W.J. Kim, T. Gould, E.K. Lee, S. Lebègue, and H. Kim. “uMBD: A materials-ready dispersion correction that uniformly treats metallic, ionic, and van der Waals bonding”. *J. Am. Chem. Soc* 142.2346 (2020).

[229] A. Ambrosetti, P.L. Silvestrelli, and A. Tkatchenko. “Physical adsorption at the nanoscale: Towards controllable scaling of the substrate-adsorbate van der Waals interaction”. *Phys. Rev. B 95* 235417 (2017).

[230] P. W. Anderson. “More Is Different”. *Science* 177.4047 (1972), pp. 393–396. DOI: `10.1126/science.177.4047.393`.

[231] John P. Perdew, Adrienn Ruzsinszky, Jianwei Sun, Niraj K. Nepal, and Aaron D. Kaplan. "Interpretations of ground-state symmetry breaking and strong correlation in wavefunction and density functional theories". *Proceedings of the National Academy of Sciences* 118.4 (2021). ISSN: 0027-8424. DOI: 10.1073/pnas.2017850118. eprint: https://www.pnas.org/content/118/4/e2017850118.full.pdf. URL: https://www.pnas.org/content/118/4/e2017850118.

[232] J. F. Janak and A. R. Williams. "Method for calculating wave functions in a nonspherical potential". *Phys. Rev. B* 23 (12 June 1981), pp. 6301–6306. DOI: 10.1103/PhysRevB.23.6301. URL: https://link.aps.org/doi/10.1103/PhysRevB.23.6301.

[233] Frank W. Kutzler and G. S. Painter. "Energies of atoms with nonspherical charge densities calculated with nonlocal density-functional theory". *Phys. Rev. Lett.* 59 (12 Sept. 1987), pp. 1285–1288. DOI: 10.1103/PhysRevLett.59.1285. URL: https://link.aps.org/doi/10.1103/PhysRevLett.59.1285.

[234] John P. Perdew, Adrienn Ruzsinszky, Lucian A. Constantin, Jianwei Sun, and Gábor I. Csonka. "Some Fundamental Issues in Ground-State Density Functional Theory: A Guide for the Perplexed". *Journal of Chemical Theory and Computation* 5.4 (2009). PMID: 26609599, pp. 902–908. DOI: 10.1021/ct800531s. eprint: https://doi.org/10.1021/ct800531s. URL: https://doi.org/10.1021/ct800531s.

[235] O. Gunnarsson and B. I. Lundqvist. "Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism". *Phys. Rev. B* 13 (10 May 1976), pp. 4274–4298. DOI: 10.1103/PhysRevB.13.4274. URL: https://link.aps.org/doi/10.1103/PhysRevB.13.4274.

[236] E. Aprà et al. "NWChem: Past, present, and future". *The Journal of Chemical Physics* 152.18 (2020), p. 184102. DOI: 10.1063/5.0004997. eprint: https://doi.org/10.1063/5.0004997. URL: https://doi.org/10.1063/5.0004997.

[237] John P. Perdew and Yue Wang. "Accurate and simple analytic representation of the electron-gas correlation energy". *Phys. Rev. B* 45 (23 June 1992), pp. 13244–13249. DOI: 10.1103/PhysRevB.45.13244. URL: https://link.aps.org/doi/10.1103/PhysRevB.45.13244.

[238] Peter W Atkins and Ronald S Friedman. *Molecular quantum mechanics*. Oxford university press, 2011.

[239] Andreas Savin. "On degeneracy, near-degeneracy and density functional theory". *Second international congress on theoretical chemical physics-ICTCP II*. 1996.

[240] George H. Booth, Deidre Cleland, Alex J. W. Thom, and Ali Alavi. “Breaking the carbon dimer: The challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods”. *The Journal of Chemical Physics* 135.8 (2011), p. 084104. DOI: 10.1063/1.3624383. eprint: https://doi.org/10.1063/1.3624383. URL: https://doi.org/10.1063/1.3624383.

[241] HA Fertig and W Kohn. “Symmetry of the atomic electron density in Hartree, Hartree-Fock, and density-functional theories”. *Physical Review A* 62.5 (2000), p. 052511.

[242] John P Perdew, Robert G Parr, Mel Levy, and Jose L Balduz Jr. “Density-functional theory for fractional particle number: derivative discontinuities of the energy”. *Physical Review Letters* 49.23 (1982), p. 1691.

[243] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. “A training algorithm for optimal margin classifiers”. *Proceedings of the fifth annual workshop on Computational learning theory*. 1992, pp. 144–152.

[244] Sebastian Dick and Marivi Fernandez-Serra. “Learning from the density to correct total energy and forces in first principle simulations”. *The Journal of chemical physics* 151.14 (2019), p. 144102.

[245] Sebastian Dick and Marivi Fernandez-Serra. “Machine learning accurate exchange and correlation functionals of the electronic density”. *Nature communications* 11.1 (2020), pp. 1–10.

[246] Xiangyun Lei and Andrew J Medford. “Design and analysis of machine learning exchange-correlation functionals via rotationally invariant convolutional descriptors”. *Physical Review Materials* 3.6 (2019), p. 063801.

[247] Jacob Hollingsworth, Li Li, Thomas E Baker, and Kieron Burke. “Can exact conditions improve machine-learned density functionals?” *The Journal of Chemical Physics* 148.24 (2018), p. 241743.

[248] Johannes Gedeon, Jonathan Schmidt, Matthew JP Hodgson, Jack Wetherell, Carlos L Benavides-Riveros, and Miguel AL Marques. “Machine learning the derivative discontinuity of density-functional theory”. *Machine Learning: Science and Technology* 3.1 (2021), p. 015011.

[249] Alexander Kolesnikov, Xiaohua Zhai, and Lucas Beyer. “Revisiting self-supervised visual representation learning”. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019, pp. 1920–1929.

[250] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. “Bert: Pre-training of deep bidirectional transformers for language understanding”. *arXiv preprint arXiv:1810.04805* (2018).

[251] Seyone Chithrananda, Gabriel Grand, and Bharath Ramsundar. “Chemberta: Large-scale self-supervised pretraining for molecular property prediction”. *arXiv preprint arXiv:2010.09885* (2020).

[252] Yu Rong, Yatao Bian, Tingyang Xu, Weiyang Xie, Ying Wei, Wenbing Huang, and Junzhou Huang. “Self-supervised graph transformer on large-scale molecular data”. *Advances in Neural Information Processing Systems* 33 (2020), pp. 12559–12571.

[253] Ashish Jaiswal, Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Debapriya Banerjee, and Fillia Makedon. “A survey on contrastive self-supervised learning”. *Technologies* 9.1 (2020), p. 2.

[254] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. “Supervised contrastive learning”. *Advances in Neural Information Processing Systems* 33 (2020), pp. 18661–18673.

[255] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. *A Simple Framework for Contrastive Learning of Visual Representations*. 2020. arXiv: 2002.05709 [cs.LG].

[256] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. “Learning long-term dependencies with gradient descent is difficult”. *IEEE transactions on neural networks* 5.2 (1994), pp. 157–166.

[257] Neena Aloysius and M Geetha. “A review on deep convolutional neural networks”. *2017 international conference on communication and signal processing (ICCSP)*. IEEE. 2017, pp. 0588–0592.

[258] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. “3D U-Net: learning dense volumetric segmentation from sparse annotation”. *International conference on medical image computing and computer-assisted intervention*. Springer. 2016, pp. 424–432.

[259] Kisuk Lee, Jonathan Zung, Peter Li, Viren Jain, and H Sebastian Seung. “Superhuman accuracy on the SNEMI3D connectomics challenge”. *arXiv preprint arXiv:1706.00120* (2017).

[260] Weiyi Gong, Tao Sun, Hexin Bai, Shah Tanvir ur Rahman Chowdhury, Peng Chu, Anoj Aryal, Jie Yu, Haibin Ling, John P. Perdew, and Qimin Yan. "Incorporation of density scaling constraint in density functional design via contrastive representation learning" (2022). URL: https://arxiv.org/abs/2205.15071.

[261] Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. "Quantum chemistry structures and properties of 134 kilo molecules". *Scientific data* 1.1 (2014), pp. 1–7.

[262] Qiming Sun, Timothy C Berkelbach, Nick S Blunt, George H Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D McClain, Elvira R Sayfutyarova, Sandeep Sharma, et al. "PySCF: the Python-based simulations of chemistry framework". *Wiley Interdisciplinary Reviews: Computational Molecular Science* 8.1 (2018), e1340.

[263] William Falcon et al. "Pytorch lightning". 3 (2019), p. 6. URL: https://github.com/PyTorchLightning/pytorch-lightning.

[264] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". *arXiv preprint arXiv:1412.6980* (2014).

[265] John P Perdew, Matthias Ernzerhof, and Kieron Burke. "Rationale for mixing exact exchange with density functional approximations". *The Journal of chemical physics* 105.22 (1996), pp. 9982–9985.

[266] G.B. Arfken, H.J. Weber, and F.E. Harris. *Bessel functions, in Mathematical Methods for Physicists*. 7th ed. Vol. 14. Academic Press, Boston, 2013, pp. 643–713.